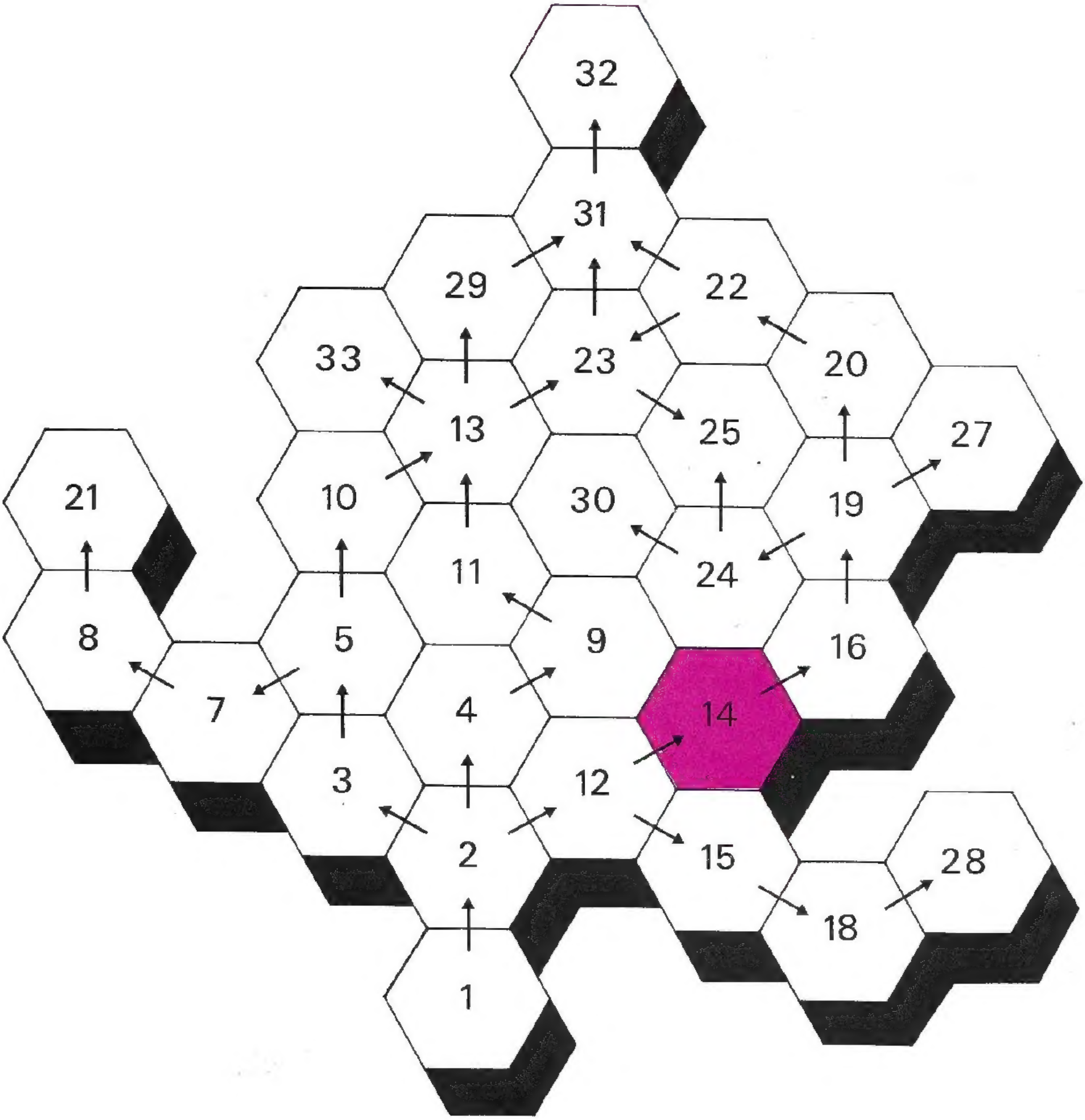
M201 14

THE OPEN UNIVERSITY

Mathematics: A Second Level Course

Linear Mathematics Unit 14

Bilinear and Quadratic Forms



The Open University

Mathematics: A Second Level Course

Linear Mathematics Unit 14

BILINEAR AND QUADRATIC FORMS

Prepared by the Linear Mathematics Course Team

The Open University Press

The Open University Press Walton Hall MK7 6AA

First published 1972. Reprinted 1976

Designed by the Media Development Group of the Open University.

Printed in Great Britain by
Martin Cadbury

SBN 335 01105 5

This text forms part of the correspondence element of an Open University Second Level Course. The complete list of units in the course is given at the end of this text.

For general availability of supporting material referred to in this text, please write to the Director of Marketing, The Open University, P.O. Box 81, Walton Hall, Milton Keynes, MK7 6AT.

Further information on Open University courses may be obtained from the Admissions Office, The Open University, P.O. Box 48, Walton Hall, Milton Keynes, MK7 6AB.

Contents

Set Books

D. L. Kreider, R. G. Kuller, D. R. Ostberg and F. W. Perkins, *An Introduction to Linear Analysis* (Addison-Wesley, 1966).

E. D. Nering, *Linear Algebra and Matrix Theory* (John Wiley, 1970).

It is essential to have these books; the course is based on them and will not make sense without them.

Conventions

Before working through this correspondence text make sure you have read *A Guide to the Linear Mathematics Course*. Of the typographical conventions given in the Guide the following are the most important.

The set books are referred to as:

K for *An Introduction to Linear Analysis*

N for *Linear Algebra and Matrix Theory*

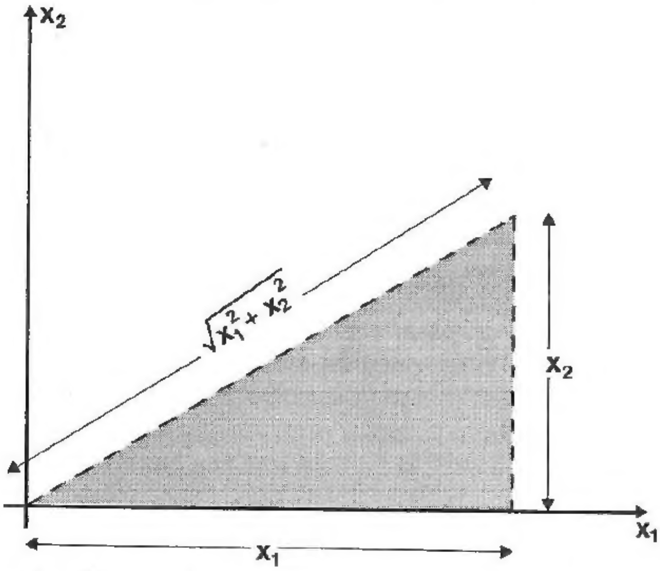
All starred items in the summaries are examinable.

References to the Open University Mathematics Foundation Course Units (The Open University Press, 1971) take the form *Unit M100 3, Operations and Morphisms.*

14.0 INTRODUCTION

This unit branches out from the single-minded pursuit of linear transformations which has occupied us in most of the other vector space units in the course so far. We shall be looking at some non-linear functions on vector spaces. An example of such a function (in the space R^2, represented by a plane) is the function that maps the number pair (x_1, x_2), represented by a point in the plane, to the square of the distance of that point from the origin; in symbols, this function is

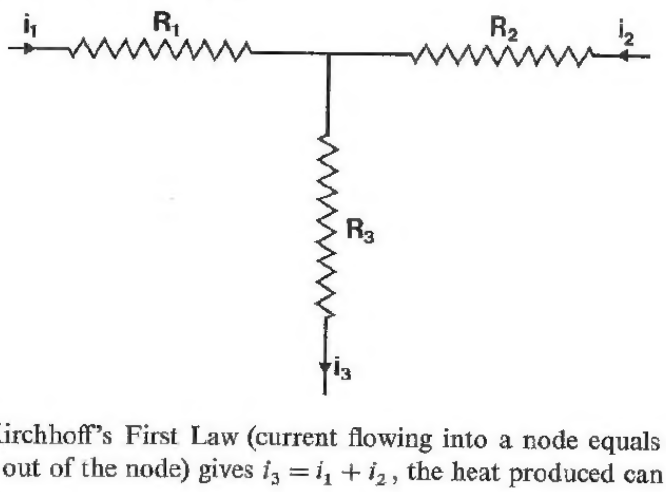
$$(x_1, x_2) \longmapsto x_1^2 + x_2^2 \qquad ((x_1, x_2) \in R^2)$$



This function is not a linear transformation; it belongs to another class of functions on vector spaces, known as *quadratic forms*, which we shall be studying later in the unit. Quadratic forms are useful in geometry not only because of their relationship to distance as described above, but also because, for instance, all the curves known as conic sections (e.g. ellipses, parabolas) can be expressed in terms of quadratic forms.

Quadratic forms also arise in various branches of applied mathematics. For example, in electric circuit theory, the heat produced in a resistor is given by Ri^2, where R is its resistance and i is the current in it; the function $i \longmapsto Ri^2$ is an example of a quadratic form. The heat produced in a more complicated circuit is given by a more complicated quadratic form; for example, in the network shown, the heat produced is

$$R_1 i_1^2 + R_2 i_2^2 + R_3 i_3^2.$$



Since Kirchhoff's First Law (current flowing into a node equals current flowing out of the node) gives $i_3 = i_1 + i_2$, the heat produced can also be written

$$R_1 i_1^2 + R_2 i_2^2 + R_3(i_1 + i_2)^2.$$

This depends on i_1 and i_2 through the function

$$(i_1, i_2) \longmapsto R_1 i_1^2 + R_2 i_2^2 + R_3(i_1 + i_2)^2 \qquad ((i_1, i_2) \in R^2),$$

which is another example of a quadratic form.

In order to relate the study of quadratic forms to linear mathematics, we approach it through the study of another type of function called a *bilinear* form.* The bilinear form associated with the quadratic form $(x_1, x_2) \longmapsto x_1^2 + x_2^2$ mentioned above is

$$((x_1, x_2), (y_1, y_2)) \longmapsto x_1 y_1 + x_2 y_2 \qquad (((x_1, x_2), (y_1, y_2)) \in R^2 \times R^2)$$

A bilinear form is a more complicated concept than a quadratic form, because its domain is not a single vector space but the Cartesian product of two vector spaces. In compensation, however, it has the advantage of being linear when we consider either of the vector spaces separately; this makes it possible to use what we already know about linear transformations to help with the study of some non-linear transformations, namely the quadratic forms. In Section 3 we obtain normal forms for matrices representing bilinear forms, and use these to analyse the stationary points of suitable differentiable real-valued functions of two variables.

* "Bilinear functional" might be a better name, because of the analogy with linear functionals.

14.1 BILINEAR FORMS

14.1.1 Definition

How much does it cost to run The Open University? Included in the total cost will be an item that depends both on the number of students taking each course and on the materials for that course: suppose that x_1 students take course number 1, x_2 take course number 2, and so on, and that y_1 is the cost of providing one correspondence package of *any* of the courses, y_2 the cost per student of one tutor-marked assignment, and y_3 the cost of one computer-marked assignment, then the item we are considering will be

$$\begin{aligned}\text{Cost} = {} & x_1(a_{11}y_1 + a_{12}y_2 + a_{13}y_3) \\ & + x_2(a_{21}y_1 + a_{22}y_2 + a_{23}y_3) \\ & + \cdots \\ & + x_n(a_{n1}y_1 + a_{n2}y_2 + a_{n3}y_3)\end{aligned}$$

where n is the number of courses, a_{r1} is the number of correspondence packages per student in the rth course, and a_{r2} the number of tutor-marked assignments and a_{r3} the number of computer-marked assignments in the rth course. The formula for cost defines a function with domain $R^n \times R^3$, since the "student numbers vector" $(x_1, \ldots, x_n)$ lies in R^n and the "cost per teaching item" vector (y_1, y_2, y_3) lies in R^3. The reason for calling this function a bilinear form is that it is linear in each of the two vectors: fix $(x_1, \ldots, x_n)$ and you get

$$(y_1, y_2, y_3) \longmapsto \text{Cost}$$

which is a linear functional; similarly, fix (y_1, y_2, y_3) and you get

$$(x_1, \ldots, x_n) \longmapsto \text{Cost}$$

which is also a linear functional.

READ from the beginning of Section IV-8 on page N156 *to the end of the second example on that page.*

Notes

(i) *Definition* The domain of f is $U \times V$, and the codomain is F. The definition of bilinearity (Equation (8.1)) can be paraphrased: for each fixed β, the function from U to F

$$\alpha \longmapsto f(\alpha, \beta) \qquad (\alpha \in U)$$

is a linear functional, and for each fixed α the function from V to F

$$\beta \longmapsto f(\alpha, \beta) \qquad (\beta \in V)$$

is a linear functional.

(ii) *Example (1)* This is a generalization of the dot product that you met in *Unit M100 22*, sub-section 1.6, in that $\{\alpha_1, \ldots, \alpha_n\}$ can be any basis, whereas the dot product which you met in the Foundation Course is expressed in terms of a basis consisting of three mutally perpendicular vectors of unit length. Many physical and mechanical laws are expressed in terms of this "geometric" dot product.

(iii) *Example (2)* Note that if we fixed α we would get the mapping

$$L_\alpha : \beta \longmapsto \int_0^1 \alpha(x)\beta(x)\,dx$$

which we discussed in *Unit 12, Linear Functionals and Duality*, sub-section 12.3.2.

Example

This example further illustrates the link between linear functionals and bilinear forms. Let V be a vector space over a field F, and consider the function f with domain $V \times \hat{V}$ and codomain F, defined by

$$f: (\alpha, \phi) \longmapsto \phi(\alpha) \qquad (\alpha \in V, \phi \in \hat{V}).$$

(Remember that $\hat{V}$ is the dual space to V.)

Fixing an element ϕ of $\hat{V}$ results in the linear functional

$$\phi : \alpha \longmapsto \phi(\alpha) \qquad (\alpha \in V)$$

on V, while fixing an element α of V results in the linear functional

$$\tilde{\alpha} : \phi \longmapsto \phi(\alpha) \qquad (\phi \in \hat{V})$$

on $\hat{V}$ which you met in sub-section 12.2.2 of *Unit 12*.
Thus f satisfies the definition of a bilinear form.

Example

If $U = V = R^3$, then the function

$$f: ((x_1, x_2, x_3), (y_1, y_2, y_3)) \longmapsto y_1(x_1 + x_2 + x_3) \qquad ((x_1, x_2, x_3), (y_1, y_2, y_3) \in R^3)$$

is a bilinear form, since by fixing y_1 we obtain a linear functional of (x_1, x_2, x_3), while by fixing (x_1, x_2, x_3) we obtain a linear functional of (y_1, y_2, y_3).

On the other hand, the function

$$g: ((x_1, x_2, x_3), (y_1, y_2, y_3)) \longmapsto x_1x_2 + y_1y_2 \qquad ((x_1, x_2, x_3), (y_1, y_2, y_3) \in R^3)$$

is ***not*** a bilinear form, since by fixing (y_1, y_2, y_3), we get the function

$$(x_1, x_2, x_3) \longmapsto x_1x_2 + \text{constant} \qquad ((x_1, x_2, x_3) \in R^3)$$

which is not a linear functional, because, for instance, if $a \in R$,

$$a(x_1, x_2, x_3) \longmapsto (ax_1)(ax_2) = a^2x_1x_2$$

and $a^2x_1x_2 \neq ax_1x_2$.

Exercise

Which of the following functions f with domain $U \times V$, codomain R, are bilinear forms?

(a) $f(a_1\alpha_1 + a_2\alpha_2, b_1\beta_1 + b_2\beta_2) = a_1 + a_2 + b_1b_2 \qquad (a_1, a_2, b_1, b_2 \in R)$, where $\{\alpha_1, \alpha_2\}$ is a basis for U and $\{\beta_1, \beta_2\}$ is a basis for V.

(b) $f(a_1\alpha_1 + a_2\alpha_2, b_1\beta_1 + b_2\beta_2) = a_1b_2 + a_2b_1$ under the same conditions as (a).

(c) $f(a_1\alpha_1 + a_2\alpha_2, b_1\beta_1 + b_2\beta_2) = a_1^2b_2^2$ under the same conditions as (a).

(d) $f(\alpha, \beta) = 0 \qquad (\alpha \in U, \beta \in V)$

(e) $f(g, h) = g'(\tfrac{1}{2})h''(\tfrac{1}{2}) \qquad (g \in U, h \in V \text{ with } U = V = C^2[0, 1])$.

Solution

The functions (b), (d) and (e) are bilinear forms. Functions (a) and (c) are not bilinear forms; this can be seen by using the argument at the end of the previous example.

In general, when testing a function from $U \times V$ into F for bilinearity, it is easier to fix each variable in turn and see if the result is a linear functional in the other variable, rather than to use Equation (8.1). For instance, function (e) in the exercise above is bilinear because, if we fix g, then $f(g, h)$ is proportional to $h''(\tfrac{1}{2})$, if we fix h, then $f(g, h)$ is proportional to $g'(\tfrac{1}{2})$, and we know that these differentiation mappings are linear.

14.1.2 Matrix Representation and Change of Basis

In our first example of a bilinear form (the one relating to the cost of The Open University), we displayed the formula in a rectangular array;

$$\begin{aligned}\text{Cost} = {} & x_1 a_{11} y_1 + x_1 a_{12} y_2 + x_1 a_{13} y_3 \\ & + x_2 a_{21} y_1 + x_2 a_{22} y_2 + x_2 a_{23} y_3 \\ & + \cdots \\ & \ \vdots \\ & + x_n a_{n1} y_1 + x_n a_{n2} y_2 + x_n a_{n3} y_3\end{aligned} \qquad (1)$$

This suggests the use of matrices, and indeed the formula can be written

$$\text{Cost} = X^T A Y$$

where $X^T = [x_1\ x_2 \cdots x_n]$ (since a one-row matrix is the transpose of a one-column matrix),

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} \end{bmatrix}$$

and

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}.$$

The matrix A, which has just the same entries as the array of coefficients in Equation (1), gives us a matrix representation of the bilinear form

$$((x_1, \ldots, x_n), (y_1, y_2, y_3)) \longmapsto \text{Cost}.$$

This use of matrices should not be confused with the use of matrices to represent linear transformations. A matrix is a device for storing information; if the information falls naturally into a rectangular array, then it can be stored as a matrix. But there are several different sorts of information that can be stored this way, and whenever a matrix is used, one must know what it is being used for. This is especially true if there is any question of changing bases in the underlying vector space(s), since the effect of a change of basis on the entries in the matrix will depend on the use to which the matrix is being put.

If U and V are m- and n-dimensional spaces respectively, then we know that a linear transformation from V to U can be represented by an $m \times n$ matrix. Presently we shall see that a *bilinear form* on $U \times V$ can also be represented by an $m \times n$ matrix, and that if P is a matrix of transition in U, Q a matrix of transition in V, then the formula for going from a matrix B representing the bilinear form in the old bases, to B' representing it in the new bases, is

$$B' = P^T B Q$$

Compare this with the formula relating two matrices (call them B and B' again) which represent a linear transformation from V to U with respect to the old and new basis (see sub-section 3.1.4 of *Unit 3, Hermite Normal Form*)

$$B' = P^{-1} B Q$$

READ from line −5 of page N156 *to line 15 of page* N158.

Note

line −7, page N157. The formula for change of basis can be derived rather more easily than via the work leading to Equations (8.4) and (8.5). We saw in *Unit 3* (page N50), that if X is a one-column matrix representing a vector in the

vector space U with respect to the old basis, and P is the matrix of transition to a new basis, then the one-column matrix X' representing the vector with respect to the new basis, is related to X by

$$X = PX'.$$

Similarly, the one column matrices Y and Y' representing a vector in the space V are related by $Y = QY'$, where Q is the matrix of transition in V.

Then, by Equation (8.3) on page N157, we have

$$\begin{aligned} f(\alpha, \beta) &= X^T B Y \\ &= (PX')^T B (QY') \\ &= (X')^T (P^T B Q) Y' \end{aligned}$$

Equation (8.3) defines the matrix representing the bilinear form in any bases, so it follows that $P^T B Q$ is the new matrix representing f.

Example of Matrix Representation

The determinant function of R^2 is a bilinear form.* As we have seen earlier (*Unit 5, Determinants and Eigenvalues*, page K681,) it is specified by

$$\det((x_1, x_2), (y_1, y_2)) = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1 y_2 - x_2 y_1,$$

and since this is equal to

$$[x_1 \; x_2] \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

the matrix representing this bilinear form is

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

with respect to the standard basis in R^2.

Example of Change of Bases

Let $U = V = R^3$, and let f be the "geometric" inner product function specified by

$$f(\xi, \eta) = x_1 y_1 + x_2 y_2 + x_3 y_3,$$

where $\xi = (x_1, x_2, x_3)$ and $\eta = (y_1, y_2, y_3)$. Then the matrix of f with respect to the standard bases in U and V is

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

because $f(\xi, \eta) = [x_1 \quad x_2 \quad x_3] B (y_1 \quad y_2 \quad y_3)$.
Now let

$$\{(1, 0, 0,), (1, 1, 0,), (1, 1, 1)\} = \{\alpha_1, \alpha_2, \alpha_3\}$$

be a new basis for U and V. The matrix of transition (whose columns represent the new basis vectors with respect to the old ones) is

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

and the matrix of f with respect to $\{\alpha_1, \alpha_2, \alpha_3\}$ is

$$B' = P^T B P = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

* The determinant function on R^n, for $n > 2$, is *not* bilinear. It is, in fact, "n-linear", or "multilinear".

Exercises

1. Exercise 1, page N159. (First see line 11, page N157)

2. Prove that congruence is an equivalence relation on matrices (see page N74).

   Hint: recall that

   (a) The identity matrix is non-singular.
   (b) The inverse of a non-singular matrix is non-singular.
   (c) The product of two non-singular matrices is non-singular.

Solutions

1. Since $x_1y_1 + 2x_1y_2 - x_2y_1 - x_2y_2 + 6x_1y_3 =$

   $$[x_1 \quad x_2]\begin{bmatrix} 1 & 2 & 6 \\ -1 & -1 & 0 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

   the matrix is

   $$\begin{bmatrix} 1 & 2 & 6 \\ -1 & -1 & 0 \end{bmatrix}.$$

2. Let "$B' \sim B$" mean "B' is congruent to B".

   We must prove three things:

   (i) $B \sim B$.
   (ii) If $B' \sim B$, then $B \sim B'$.
   (iii) If $B' \sim B$ and $B'' \sim B'$, then $B'' \sim B$.

   We prove (i) by noting that $B = I^TBI$, and I is non-singular.

   We prove (ii) by noting that if $B' \sim B$, then $B' = P^TBP$, where P is non-singular. Thus

   $$\begin{aligned}(P^{-1})^TB'P^{-1} &= (P^{-1})^TP^TBPP^{-1} \\ &= (P^T)^{-1}P^TBPP^{-1} \\ &= IBI \\ &= B\end{aligned}$$

   and P^{-1} is non-singular. Thus $B \sim B'$.

   We prove (iii) by noting that, if $B' = P^TBP$ and $B'' = Q^TB'Q$, where P, Q are non-singular, then

   $$\begin{aligned}B'' &= Q^T(P^TBP)Q \\ &= (PQ)^TB(PQ)\end{aligned}$$

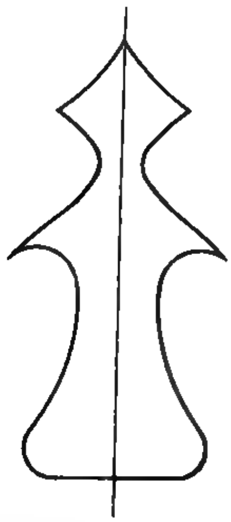
   where PQ is non-singular.

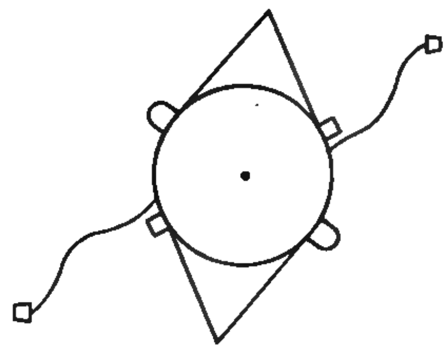
   Thus

   $$B'' \sim B.$$

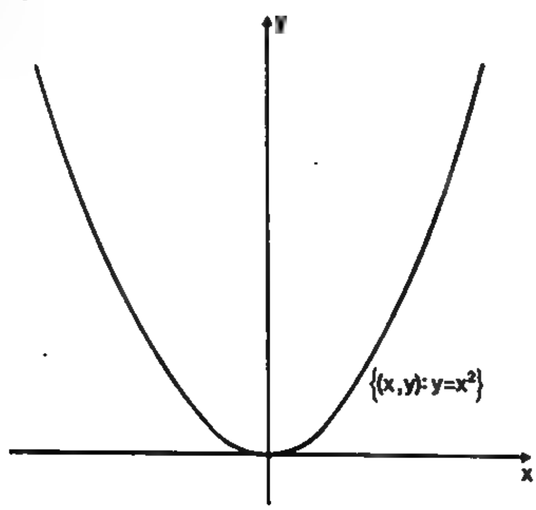
14.1.3 Symmetric and Skew-symmetric Forms

The next idea we consider is that a bilinear form can be *symmetric*. What does this mean? We saw in *Unit M100 30, Groups I*, that objects can have various sorts of symmetry, depending on what sort of transformations leave them invariant. For example, the shape in the left-hand figure is symmetric with respect to a reflection about the vertical axis shown, and the shape in the right-hand figure is symmetric with respect to a rotation of 180° about the point shown.





The graph of the function $f: x \longmapsto x^2$ is symmetric with respect to reflection in the y-axis.



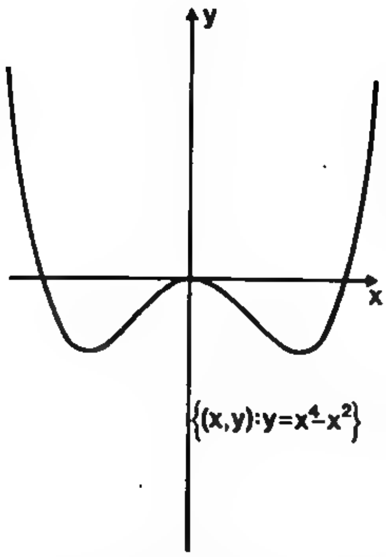
In this case the symmetry can be expressed algebraically by the equation:

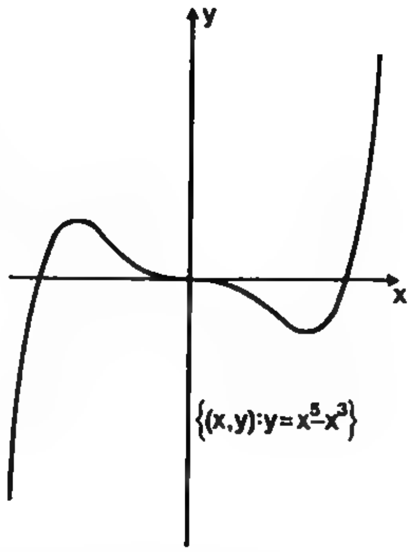
$$f(x) = f(-x) \qquad (x \in R).$$

Functions with this property are said to be *even*. For example, the cosine function is even. This sort of symmetry is interesting from our point of view, in that we can form a complementary concept, that of *anti-symmetry*, expressed by the equation

$$g(x) = -g(-x) \qquad (x \in R).$$

Functions with this property are said to be *odd*. For example, the sine function and the function $g: x \longmapsto x^3$ are anti-symmetric or odd. In fact, any polynomial function all of whose non-zero terms are in even powers of x, is symmetric, and any polynomial function all of whose non-zero terms are in odd powers of x, is anti-symmetric.





What do we know about an *arbitrary* polynomial function of x? Precisely that it is the *sum* of its odd terms and its even terms. For instance,

$$h: x \longmapsto 1 - x + 2x^2 - 4x^5 \qquad (x \in R)$$

is equal to $f + g$, where

$$f: x \longmapsto 1 + 2x^2$$

is even and

$$g: x \longmapsto -x - 4x^5$$

is odd. This is not at all difficult to see: what is interesting is that *any* function $h: R \longrightarrow R$ can be expressed as the sum of a symmetric part, f, and an anti-symmetric part, g. For, any combination

$$x \longmapsto ah(x) + ah(-x) \qquad a \in R,$$

is symmetric, and any combination

$$x \longmapsto bh(x) - bh(-x) \qquad b \in R,$$

is anti-symmetric. The expressions for f and g are found by choosing a and b so that $f + g = h$, i.e.

$$f(x) = \tfrac{1}{2}(h(x) + h(-x)) \qquad (x \in R)$$
$$g(x) = \tfrac{1}{2}(h(x) - h(-x)) \qquad (x \in R).$$

Example

The function $x \longmapsto e^x$ can be expressed as the sum of the symmetric part

$$\cosh: x \longmapsto \tfrac{1}{2}(e^x + e^{-x})$$

and the anti-symmetric part

$$\sinh: x \longmapsto \tfrac{1}{2}(e^x - e^{-x}).$$

The way of splitting h into symmetric and anti-symmetric parts is unique; however it is done, the same f and g are obtained.

Exercise

Prove that if $h = k + l$, where k is symmetric and l is anti-symmetric, then $k = f$ and $l = g$. (The functions, f, g, h are as above.)

Solution

$$h(x) = k(x) + l(x) \qquad (x \in R) \qquad (1)$$

so that

$$h(-x) = k(-x) + l(-x) \qquad (x \in R)$$
$$= k(x) - l(x) \qquad (2)$$

because k is symmetric and l is anti-symmetric. Solving Equations (1) and (2) for k and l, we find that $k = f$ and $l = g$.

We now consider symmetry for bilinear forms. Since

$$f(\alpha, \beta) = f(-\alpha, -\beta)$$

is a tautology (f is linear in α and β separately), the obvious extension of the definition of symmetry for functions of one variable gets us nowhere. The definitions we consider are not directly analogous to those discussed above for functions of one variable, but the results are very similar. The definitions involve interchanging the two vectors in the ordered pair (α, β).

Definitions

A bilinear form $f: U \times U \longrightarrow F$, is *symmetric* if $f(\alpha, \beta) = f(\beta, \alpha)$ for all $\alpha, \beta \in U$.

A bilinear form $g: U \times U \longrightarrow F$, is *anti-symmetric* if $g(\alpha, \beta) = -g(\beta, \alpha)$ for all $\alpha, \beta \in U$.

Examples

The determinant function on R^2

$$((x_1, x_2), (y_1, y_2)) \longmapsto x_1 y_2 - x_2 y_1$$

is antisymmetric.

The function

$$((x_1, x_2), (y_1, y_2)) \longmapsto x_1 y_2 + x_2 y_1$$

with domain R^2, however, is symmetric.

Exercise

Find expressions for the symmetric and anti-symmetric parts of a general bilinear form $h: U \times U \longrightarrow F$.

(*Hint*: the method is similar to that for real functions.)

Solution

The symmetric part is

$$f: (\alpha, \beta) \longmapsto \tfrac{1}{2}(h(\alpha, \beta) + h(\beta, \alpha)) \qquad (\alpha, \beta \in U)$$

and the anti-symmetric part is

$$g: (\alpha, \beta) \longmapsto \tfrac{1}{2}(h(\alpha, \beta) - h(\beta, \alpha)) \qquad (\alpha, \beta \in U)$$

But wait! This is all right if $F = R$, but for a general field F, what do we mean by $\frac{1}{2}$? What we mean, of course, is the (multiplicative) inverse of the element $1 + 1$, which exists as long as $1 + 1 \neq 0$, the additive identity element. But for a general field F, we have no guarantee that $1 + 1 \neq 0$; and in fact there is one field where $1 + 1 = 0$, namely the field consisting of the two elements 0 and 1, with addition and multiplication tables:

+	0	1
0	0	1
1	1	0

×	0	1
0	0	0
1	0	1

In this field there is one other thing to settle: what do we mean by the minus in, say

$$g(\alpha, \beta) = -g(\beta, \alpha)?$$

If we write it in the form $g(\alpha, \beta) + g(\beta, \alpha) = 0$, then there is no problem, and we recognize that $-g(\beta, \alpha)$ is the (additive) inverse of $g(\alpha, \beta)$. In our present field we have only two elements and

$$0 + 0 = 0, \quad 1 + 1 = 0$$

so that $0 = -0$, $1 = -1$. Hence our definition for anti-symmetry can be written

$$g(\alpha, \beta) = g(\beta, \alpha),$$

i.e. it is the same as symmetry. We can therefore conclude that, for this sort of field, a bilinear form is symmetric if and only if it is anti-symmetric!

Because of this anomaly, (and it isn't any more than that; in general we shall ignore it) N defines *skew-symmetry*, which is the same as anti-symmetry when $1 + 1 \neq 0$, but provides additional information about the bilinear form in the case in which $1 + 1 = 0$.

Definition

A bilinear form $g: U \times U \longrightarrow F$ is *skew-symmetric* if $g(\alpha, \alpha) = 0$ for all $\alpha \in U$.

If a bilinear form is skew-symmetric, then it is always anti-symmetric; for the definition of bilinear form implies, for all $\alpha, \beta \in U$, that

$$g(\alpha + \beta, \alpha + \beta) = g(\alpha, \alpha) + g(\alpha, \beta) + g(\beta, \alpha) + g(\beta, \beta),$$

and the definition of skew-symmetric implies

$$g(\alpha + \beta, \alpha + \beta) = g(\alpha, \alpha) = g(\beta, \beta) = 0;$$

so the two together give

$$g(\alpha, \beta) + g(\beta, \alpha) = 0$$

or equivalently, $g(\alpha, \beta) = -g(\beta, \alpha)$.

On the other hand, the converse only works if $1 + 1 \neq 0$. If g is anti-symmetric, then taking the general equation

$$g(\alpha, \beta) = -g(\beta, \alpha)$$

and letting $\alpha = \beta$, we get

$$g(\alpha, \alpha) = -g(\alpha, \alpha);$$
$$g(\alpha, \alpha) + g(\alpha, \alpha) = 0;$$
$$(1 + 1)g(\alpha, \alpha) = 0;$$

so that, if $1 + 1 \neq 0$, we deduce $g(\alpha, \alpha) = 0$ and the form is skew-symmetric.

Thus, as long as $1 + 1 \neq 0$, there is no difference between the concepts of anti-symmetry and skew-symmetry. As we deal exclusively with R or C as the fields in all applications in this course, you need only remember one of the definitions.

READ from line 16 on page N158 *to the end of Section IV-8 on page* N159.

Notes

(i) It is important to note the "if and only if" in the statement of ***Theorem 8.1***, and the corresponding two parts of the proof. Also notice that in the skew-symmetric case there is no "if and only if" theorem because of the possibility of $1 + 1 = 0$. Hence ***Theorems 8.2*** and ***8.3***.

(ii) Note the definitions of symmetry and skew-symmetry of ***matrices*** that occur in the course of the proof of ***Theorem 8.1***, and immediately after the proof of ***Theorem 8.3***. Of course, a non-square matrix cannot be symmetric or skew-symmetric; nor can a bilinear form on $U \times V$ if $V \neq U$.

Exercises

1. Exercise 2, page N159.
2. Exercise 3, page N159.
3. Exercise 4, page N159.

Solutions

1. The symmetric part is

$$\frac{1}{2}\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\\7 & 8 & 9\end{bmatrix}+\frac{1}{2}\begin{bmatrix}1 & 4 & 7\\2 & 5 & 8\\3 & 6 & 9\end{bmatrix}=\begin{bmatrix}1 & 3 & 5\\3 & 5 & 7\\5 & 7 & 9\end{bmatrix}$$

The skew-symmetric part is

$$\frac{1}{2}\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\\7 & 8 & 9\end{bmatrix}-\frac{1}{2}\begin{bmatrix}1 & 4 & 7\\2 & 5 & 8\\3 & 6 & 9\end{bmatrix}=\begin{bmatrix}0 & -1 & -2\\1 & 0 & -1\\2 & 1 & 0\end{bmatrix}$$

2. The transpose of P^TBP is the product of the transposes of P^T, B and P (see *Unit 3*, sub-section 3.2.4), reversed in order, i.e.

$$(P^TBP)^T=(P)^TB^T(P^T)^T=P^TB^TP.$$

Thus, if B is symmetric, then $B^T=B$ and so $(P^TBP)^T=P^TBP$, i.e. P^TBP is symmetric. If B is skew-symmetric, then $B^T=-B$, so that

$$(P^TBP)^T=P^T(-B)P=-(P^TBP),$$

i.e. P^TBP is skew-symmetric.

3. If A is an $m \times n$ matrix, then A^TA is the product of an $n \times m$ and an $m \times n$ matrix. This product is defined, and is $n \times n$. Similarly, the product AA^T is defined, and is $m \times m$. Furthermore, $(A^TA)^T=(A)^T(A^T)^T=A^TA$, so A^TA is symmetric. Similarly, AA^T is symmetric.

14.1.4 Summary of Section 14.1

In this section we defined the terms

bilinear form	(page N156)	* * *
symmetric bilinear form	(page N158)	* * *
anti-symmetric bilinear form	(page C14)	*
skew-symmetric bilinear form	(page N158)	* * *
symmetric matrix	(page N158)	* * *
skew-symmetric matrix	(page N159)	* * *
congruent	(page N158)	* * *

Theorems

1. (8.1, page N158)
A bilinear form f is symmetric if and only if any matrix B representing f has the property $B^T=B$. * * *

2. (8.2, page N158)
If a bilinear form f is skew-symmetric, then any matrix B representing f has the property $B^T=-B$. * * *

3. (8.3, page N158)
If $1+1\neq 0$ and the matrix B representing f has the property $B^T=-B$, then f is skew-symmetric. * * *

4. (8.4, page N159)
If $1+1\neq 0$, every bilinear form can be represented uniquely as a sum of a symmetric bilinear form and a skew-symmetric bilinear form. * * *

Technique

Given a bilinear form f, find f_s, f_{ss}. * * *

Notation

$f_s(\alpha,\beta)$ (page N159)
$f_{ss}(\alpha,\beta)$ (page N159)

14.2 QUADRATIC FORMS

14.2.1 Definition

We are now in a position to look at certain *non-linear* functions from a vector space to its field, and analyse these in terms of bilinear forms, thus giving us the opportunity to use linear tools.

Definition

A quadratic form on a vector space V is a function $q: V \longrightarrow F$, such that

$$q(\alpha) = f(\alpha, \alpha) \qquad (\alpha \in V)$$

for some bilinear form f on $V \times V$. (Throughout the rest of this unit, the bilinear forms are always on $V \times V$ rather than $U \times V$; so we can compress the notation and call them bilinear forms on V.)

A quadratic form is not a linear form (except in the trivial case where it is the zero mapping); the following are among the simplest examples of quadratic forms.

Examples

1. The function q, where

   $$q(x) = x^2 \qquad (x \in R)$$

   is a quadratic form on R, obtained from the bilinear form f, where

   $$f(x, y) = xy \qquad (x, y \in R)$$

   since $f(x, x) = x^2$.

2. The function Q, where

   $$Q(x_1, x_2) = x_1^2 + 2x_1x_2 - x_2^2 \qquad ((x_1, x_2) \in R^2)$$

   is a quadratic form, since it can be obtained from the bilinear form F, where

   $$F((x_1, x_2), (y_1, y_2)) = x_1y_1 + 2x_1y_2 - x_2y_2 \qquad ((x_1, x_2), (y_1, y_2) \in R^2).$$

To discover the quadratic form q corresponding to a given bilinear form f, we simply use the equation $q(\alpha) = f(\alpha, \alpha)$. For example, if

$$f((x_1, x_2), (y_1, y_2)) = 2x_1y_1 + 3x_2y_2 \qquad ((x_1, x_2), (y_1, y_2) \in R^2),$$

then the corresponding quadratic form is specified by

$$\begin{aligned} q(x_1, x_2) &= f((x_1, x_2), (x_1, x_2)) \\ &= 2x_1x_1 + 3x_2x_2 \\ &= 2x_1^2 + 3x_2^2 \qquad ((x_1, x_2) \in R^2) \end{aligned}$$

Similarly, if f is the bilinear form on $C[0, 1]$ specified by

$$f(g, h) = \int_0^1 g(x)h(x)\,dx \qquad (g, h \in C[0, 1]),$$

then the corresponding quadratic form is specified by setting $g = h$.

$$\begin{aligned} q(g) &= f(g, g) \\ &= \int_0^1 [g(x)]^2\,dx \qquad (g \in C[0, 1]). \end{aligned}$$

Exercises

1. Write down the quadratic forms corresponding to the following bilinear forms*:

   (i) $x_1y_1 + x_2y_2$ $\quad ((x_1, x_2), (y_1, y_2) \in R^2)$

   (ii) $x_1y_2 + x_2y_1$ $\quad ((x_1, x_2), (y_1, y_2) \in R^2)$

   (iii) $2x_1y_2$ $\quad ((x_1, x_2), (y_1, y_2) \in R^2)$

   (iv) $\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}$ $\quad ((x_1, x_2), (y_1, y_2) \in R^2)$

   (v) $\int_0^1 f(x)g(1-x)\,dx$ $\quad (f, g \in C[0, 1])$

   (vi) $\int_0^1 x^2 f(x)g(x)\,dx$ $\quad (f, g \in C[0, 1])$

2. Let q be any quadratic form on a vector space V, whose field is R. Show that, for any $a \in R$ and any vector $\alpha \in V$,

$$q(a\alpha) = a^2q(\alpha).$$

Solutions

1. In abbreviated form the quadratic forms are:

   (i) $x_1^2 + x_2^2$ $\quad ((x_1, x_2) \in R^2)$

   (ii) $2x_1x_2$ $\quad ((x_1, x_2) \in R^2)$

   (iii) $2x_1x_2$ $\quad ((x_1, x_2) \in R^2)$

   (iv) $\begin{vmatrix} x_1 & x_1 \\ x_2 & x_2 \end{vmatrix} (= 0)$ $\quad ((x_1, x_2) \in R^2)$

   (v) $\int_0^1 f(x)f(1-x)\,dx$ $\quad (f \in C[0, 1])$

   (vi) $\int_0^1 x^2[f(x)]^2\,dx$ $\quad (f \in C[0, 1])$

2. Since q is a quadratic form, we have

$$\begin{aligned} q(a\alpha) &= f(a\alpha, a\alpha) && \text{where } f \text{ is bilinear} \\ &\left.\begin{aligned} &= af(\alpha, a\alpha) \\ &= a^2f(\alpha, \alpha) \end{aligned}\right\} && \text{since } f \text{ is a bilinear form} \\ &= a^2q(\alpha). \end{aligned}$$

14.2.2 The Polar Form of a Quadratic Form

In parts (ii) and (iii) of Solution 1 of the previous sub-section, we saw that two different bilinear forms, $x_1y_2 + x_2y_1$ and $2x_1y_2$, give rise to the same quadratic form $2x_1x_2$. However, only the first of the bilinear forms is symmetric. The next reading passage shows that this is a special case of a general result:

> any given quadratic form q can be obtained from many different bilinear forms f, using the formula
>
> $$q(\alpha) = f(\alpha, \alpha)$$
>
> but only one of the bilinear forms is symmetric.

READ Section IV-9 on page N160 *to the end on page* N162.

* We have abbreviated the notation, writing, for example

$$x_1y_1 \quad (x_1, y_1 \in R)$$

in place of

$$(x_1, y_1) \longmapsto x_1y_1 \quad (x_1, y_1 \in R).$$

This practice is adopted by N.

Notes

(i) *Paragraph preceding Theorem 9.1, page* N161
Note the suggestive mnemonic:

$$xy = \tfrac{1}{2}[(x+y)^2 - x^2 - y^2].$$

This gives a helpful way of remembering the formula

$$f_s(\alpha, \beta) = \tfrac{1}{2}[q(\alpha+\beta) - q(\alpha) - q(\beta)].$$

(ii) *line −10, page* N161 By the matrix representing a quadratic form q we mean a matrix A such that

$$q(\alpha) = X^T A X \qquad (\alpha \in V)$$

where X is the one-column matrix representing α. This matrix also represents the bilinear form f defined by

$$f(\alpha, \beta) = X^T A Y \qquad (\alpha, \beta \in V)$$

where Y is the one-column matrix representing β. This bilinear form is one of those with the property

$$q(\alpha) = f(\alpha, \alpha) \qquad (\alpha \in V)$$

and if A is symmetric f is the polar form of q.
(iii) *line −6, page* N161 *to the end of the section* The details of the geometrical interpretation are not important. The geometry involved is not part of the course.

Example

If q is the quadratic form specified by

$$q(x_1, x_2) = x_1^2 + 4x_1x_2 + 2x_2^2 \qquad ((x_1, x_2) \in R^2),$$

then in matrix notation,

$$q(x_1, x_2) = [x_1\ x_2]\begin{bmatrix}1 & 4\\ 0 & 2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} = X^T A X$$

so A is a matrix representing q. A also represents the bilinear form specified by

$$f((x_1, x_2), (y_1, y_2)) = x_1y_1 + 4x_1y_2 + 2x_2y_2,$$

which is a bilinear form (not symmetric, however) such that $q(\alpha) = f(\alpha, \alpha)$. On the other hand, the *symmetric* bilinear form corresponding to q, i.e. the *polar* form, is by Equation (9.1) on page N161

$$f_s((x_1, x_2), (y_1, y_2)) = x_1y_1 + 2x_1y_2 + 2x_2y_1 + 2x_2y_2$$

which has the matrix

$$A_s = \begin{bmatrix}1 & 2\\ 2 & 2\end{bmatrix}$$

and it is still true that

$$q(x_1, x_2) = X^T A_s X.$$

Exercise

Write down the polar form of each of the following quadratic forms.

(i) x_1^2 $\qquad (x_1 \in R)$

(ii) $x_1^2 + x_2^2$ $\qquad ((x_1, x_2) \in R^2)$

(iii) $x_1^2 + 2x_1x_2 + 3x_2x_3 + x_3^2$ $\qquad ((x_1, x_2, x_3) \in R^3)$

(iv) $\int_0^1 f(x)f(1-x)\,dx$ $\qquad (f \in C[0, 1])$.

Solution

(i) $$f((x_1, y_1)) = \tfrac{1}{2}[q(x_1 + y_1) - q(x_1) - q(y_1)]$$
$$= \tfrac{1}{2}[(x_1 + y_1)^2 - x_1^2 - y_1^2]$$
$$= x_1 y_1$$

(ii) $$f((x_1, x_2), (y_1, y_2)) = \tfrac{1}{2}[q(x_1 + y_1, x_2 + y_2) - q(x_1, x_2) - q(y_1, y_2)]$$
$$= \tfrac{1}{2}[(x_1 + y_1)^2 + (x_2 + y_2)^2 - x_1^2 - x_2^2 - y_1^2 - y_2^2]$$
$$= x_1 y_1 + x_2 y_2$$

(iii) $$f((x_1, x_2, x_3), (y_1, y_2, y_3))$$
$$= \tfrac{1}{2}[q(x_1 + y_1, x_2 + y_2, x_3 + y_3) - q(x_1, x_2, x_3) - q(y_1, y_2, y_3)]$$
$$= \tfrac{1}{2}[(x_1 + y_1)^2 + 2(x_1 + y_1)(x_2 + y_2) + 3(x_2 + y_2)(x_3 + y_3) + (x_3 + y_3)^2 - x_1^2 - 2x_1x_2 - 3x_2x_3 - x_3^2 - y_1^2 - 2y_1y_2 - 3y_2y_3 - y_3^2]$$
$$= x_1y_1 + x_1y_2 + x_2y_1 + \tfrac{3}{2}x_2y_3 + \tfrac{3}{2}x_3y_2 + x_3y_3.$$

(iv) If we let Q be the quadratic form given, and F the polar form of Q, then, for $f, g \in C[0, 1]$:

$$F(f, g) = \tfrac{1}{2}[Q(f + g) - Q(f) - Q(g)]$$
$$= \tfrac{1}{2}\left\{\int_0^1 [f(x) + g(x)][f(1 - x) + g(1 - x)]\, dx - \int_0^1 f(x)f(1 - x)\, dx - \int_0^1 g(x)g(1 - x)\, dx\right\}$$
$$= \tfrac{1}{2}\int_0^1 [f(x)g(1 - x) + g(x)f(1 - x)]\, dx.$$

In general, the polar form is not the simplest bilinear form corresponding to a given quadratic form. For example, in part (iv) of the last exercise, a simpler bilinear form is

$$H(f, g) = \int_0^1 f(x)g(1 - x)\, dx$$

and we still have

$$Q(f) = H(f, f).$$

You may have noticed that in parts (i), (ii) and (iii) of the exercise there is a rule of thumb which gives the answer considerably more speedily than by applying the formula

$$f(\alpha, \beta) = \tfrac{1}{2}[q(\alpha + \beta) - q(\alpha) - q(\beta)].$$

This rule is: every term of the type x_i^2 in the quadratic form becomes $x_i y_i$ in the polar form, whereas each term of the type $x_i x_j$ $(i \neq j)$ in the quadratic form becomes $\tfrac{1}{2}(x_i y_j + x_j y_i)$ in the polar form. This rule always works for quadratic forms which are expressed in terms of coordinates with respect to a basis. However, to do (iv) above without guesswork, one needs a formula which is not expressed in terms of a basis.

Exercises

1. Write down the polar forms of each of the following quadratic forms, using the above rule.

   (i) $x_1x_2 + x_2x_3 \qquad ((x_1, x_2, x_3) \in R^3)$

   (ii) $x_1^2 + x_1x_2 + x_3^2 \qquad ((x_1, x_2, x_3) \in R^3)$

   (iii) $-x_1^2 + x_2x_3 \qquad ((x_1, x_2, x_3) \in R^3)$

   (iv) $x_1^2 + \frac{1}{2}x_1x_2 - x_2^2 + \frac{3}{2}x_2x_3 + 2x_3^2 \qquad ((x_1, x_2, x_3) \in R^3)$

2. Write down the symmetric matrices representing each of the quadratic forms of Exercise 1.

Solutions

1. ((x_1, x_2, x_3) and $(y_1, y_2, y_3) \in R^3$ throughout.)

   (i) $\frac{1}{2}(x_1y_2 + x_2y_1 + x_2y_3 + x_3y_2)$

   (ii) $x_1y_1 + \frac{1}{2}(x_1y_2 + x_2y_1) + x_3y_3$

   (iii) $-x_1y_1 + \frac{1}{2}(x_2y_3 + x_3y_2)$

   (iv) $x_1y_1 + \frac{1}{4}(x_1y_2 + x_2y_1) - x_2y_2 + \frac{3}{4}(x_2y_3 + x_3y_2) + 2x_3y_3$

2.

   (i) $\begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

   (iii) $\begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{bmatrix}$ (iv) $\begin{bmatrix} 1 & \frac{1}{4} & 0 \\ \frac{1}{4} & -1 & \frac{3}{4} \\ 0 & \frac{3}{4} & 2 \end{bmatrix}$

14.2.3 The Quadratic Taylor Approximation

In this sub-section, we will apply quadratic forms to the problem of approximating complicated functions by simpler ones. In sub-section 14.3.4, we use such approximations to classify the stationary points of functions of two variables.

In *Unit M100 14, Sequences and Limits II*, we saw that a function f of one real variable can be approximated by polynomials, provided it and its derivatives satisfy suitable conditions. The higher the degree of the polynomial, the better the approximation. These approximations are known as *Taylor approximations*, and are obtained by calculating the derivatives of the function at a suitable point. If we choose the point x_0, then the nth-degree Taylor approximation to $f(x)$ in the neighbourhood of x_0 is given by:

$$f(x_0) + (x - x_0)f'(x_0) + \frac{(x - x_0)^2}{2}f''(x_0) + \cdots + \frac{(x - x_0)^n}{n!}f^{(n)}(x_0).$$

Now there is a similar approximation process for functions of more than one real variable (i.e. for functions with domain $R^2, R^3, \ldots$). The formulas in this case are expressed in terms of *partial derivatives*, which you met in *Unit M100 15, Differentiation II*, for functions of two real variables. We give a brief résumé of the concept of a partial derivative below.

If f is a function with domain R^2 and codomain R, let us for convenience call the first variable x_1 and the second variable x_2. For instance,

$$f: (x_1, x_2) \longmapsto e^{x_1} + x_1 \sin x_2$$

is a function of the sort we are interested in.

Then the *first partial derivatives* are obtained by fixing one of the variables and considering the derivative of f considered as a function of the other variable only. In symbols*

$$f_1(x_1, x_2) = \lim_{h\to 0} \frac{f(x_1 + h, x_2) - f(x_1, x_2)}{h}$$

$$f_2(x_1, x_2) = \lim_{h\to 0} \frac{f(x_1, x_2 + h) - f(x_1, x_2)}{h}$$

Second partial derivatives are similarly defined: $f_{11} = (f_1)_1, f_{12} = (f_1)_2$, etc. For instance, in the case of f above:

$$f_1(x_1, x_2) = e^{x_1} + \sin x_2$$

(since all we have to do is to differentiate with respect to x_1, regarding x_2 as a constant).

Similarly:

$$f_2(x_1, x_2) = x_1 \cos x_2$$

The partial derivative of f_1 with respect to x_1, is

$$f_{11}(x_1, x_2) = e^{x_1}.$$

Also,

$$f_{12}(x_1, x_2) = \cos x_2$$
$$f_{21}(x_1, x_2) = \cos x_2$$
$$f_{22}(x_1, x_2) = -x_1 \sin x_2.$$

Notice that in this case, $f_{12} = f_{21}$. This will always be true in the cases which interest us. (In fact, the condition that all the second partial derived functions exist *and are continuous*, is enough to guarantee that $f_{12} = f_{21}$.)

Exercises

1. Find all the first and second partial derivatives of the following functions, where the domain is R^2 in each case.

   (i) $f: (x_1, x_2) \longmapsto 3x_1x_2 + 5x_2^2 + 3x_1^3x_2$.
   (ii) $g: (x_1, x_2) \longmapsto x_1 \exp(x_1 + x_2)$
   (iii) $h: (x_1, x_2) \longmapsto x_1 \cos(2x_1 + x_2)$

2. Find $h_1(0, 0)$ and $h_{12}\left(\frac{\pi}{4}, 0\right)$ where h is as in part (iii) of Exercise 1.

Solutions

1. (i) $f_1(x_1, x_2) = 3x_2 + 9x_1^2x_2$
   $f_2(x_1, x_2) = 3x_1 + 10x_2 + 3x_1^3$
   $f_{11}(x_1, x_2) = 18x_1x_2$
   $f_{12}(x_1, x_2) = f_{21}(x_1, x_2) = 3 + 9x_1^2$
   $f_{22}(x_1, x_2) = 10$

   (ii) $g_1(x_1, x_2) = (x_1 + 1) \exp(x_1 + x_2)$
   $g_2(x_1, x_2) = x_1 \exp(x_1 + x_2)$
   $g_{11}(x_1, x_2) = (x_1 + 2) \exp(x_1 + x_2)$
   $g_{12}(x_1, x_2) = g_{21}(x_1, x_2) = (x_1 + 1) \exp(x_1 + x_2)$
   $g_{22}(x_1, x_2) = x_1 \exp(x_1 + x_2)$

* Note that we have dropped the prime from f_1', f_2', the forms used in the Foundation Course.

(iii) $h_1(x_1, x_2) = \cos(2x_1 + x_2) - 2x_1 \sin(2x_1 + x_2)$

$$h_2(x_1, x_2) = -x_1 \sin(2x_1 + x_2)$$
$$h_{11}(x_1, x_2) = -4\sin(2x_1 + x_2) - 4x_1 \cos(2x_1 + x_2)$$
$$h_{12}(x_1, x_2) = h_{21}(x_1, x_2)$$
$$= -\sin(2x_1 + x_2) - 2x_1 \cos(2x_1 + x_2)$$
$$h_{22}(x_1, x_2) = -x_1 \cos(2x_1 + x_2)$$

2. $h_1(0, 0) = \cos(0) - 2 \times 0 \times \sin(0)$
$= 1$

$$h_{12}\left(\frac{\pi}{4}, 0\right) = -\sin\left(\frac{\pi}{2}\right) - \frac{\pi}{2}\cos\left(\frac{\pi}{2}\right)$$
$$= -1.$$

We are now ready to look at the Taylor approximation method in two variables. The most general polynomial function of first degree in two variables has the form

$$p: (x_1, x_2) \longmapsto a + bx_1 + cx_2 \qquad ((x_1, x_2) \in R^2)$$

and the first-degree Taylor approximation to f will be a first-degree polynomial function of this sort. The approximation about (0, 0), say, will be obtained by requiring the image of p and its first partial derivatives at (0, 0) to agree with the image of f and its first partial derivatives at (0, 0). These three conditions give just enough information to determine the three numbers a, b, c:

$$f(0, 0) = a,$$
$$f_1(0, 0) = b,$$
$$f_2(0, 0) = c.$$

Exercise

Show that the above choice of a, b and c makes the image of the function and its first partial derivatives at (0, 0) agree with the image of the approximation and its first partial derivatives at (0, 0).

Solution

The approximation is the polynomial function p specified by

$$p(x_1, x_2) = f(0, 0) + x_1 f_1(0, 0) + x_2 f_2(0, 0) \qquad ((x_1, x_2) \in R^2).$$

Thus
$$p_1(x_1, x_2) = f_1(0, 0) \qquad ((x_1, x_2) \in R^2)$$
$$p_2(x_1, x_2) = f_2(0, 0) \qquad ((x_1, x_2) \in R^2).$$

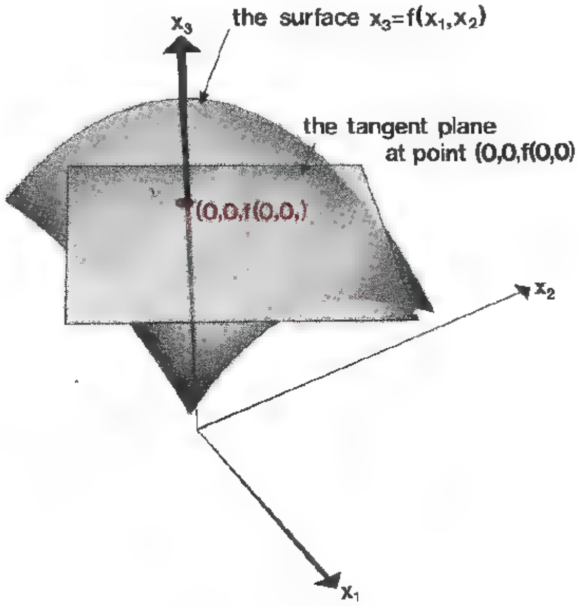
Then, setting $(x_1, x_2) = (0, 0)$ in the above three equations, we get

$$p(0, 0) = f(0, 0),$$
$$p_1(0, 0) = f_1(0, 0),$$
$$p_2(0, 0) = f_2(0, 0).$$

Thus the first-degree Taylor approximation is

$$f(x_1, x_2) \simeq f(0, 0) + x_1 f_1(0, 0) + x_2 f_2(0, 0).$$

Geometrically, this corresponds to approximating the three-dimensional surface $x_3 = f(x_1, x_2)$ by its tangent plane at the point $(0, 0, f(0, 0))$, in the manner explained in *Unit M100 15*.



The same method can be extended to give higher-degree Taylor approximations. For instance the second-degree Taylor approximation to f about $(0, 0)$ can be obtained by writing out the most general second-degree polynomial in x_1, x_2:

$$r(x_1, x_2) = A + Bx_1 + Cx_2 + Dx_1^2 + Ex_1x_2 + Fx_2^2$$

We now choose the numbers $A, B, \ldots, F$ to make the images of r and f agree in their values at $(0, 0)$, and in the values of the *first and second* partial derivatives, all at $(0, 0)$.

Exercise

Verify that the appropriate values to take are:

$$A = f(0, 0), \quad B = f_1(0, 0), \quad C = f_2(0, 0), \quad D = \tfrac{1}{2}f_{11}(0, 0),$$
$$E = f_{12}(0, 0), \quad F = \tfrac{1}{2}f_{22}(0, 0).$$

Solution

With these values, the approximation is given by:

$$\begin{aligned} r(x_1, x_2) = {} & f(0, 0) + x_1 f_1(0, 0) + x_2 f_2(0, 0) \\ & + \tfrac{1}{2}x_1^2 f_{11}(0, 0) + x_1x_2 f_{12}(0, 0) + \tfrac{1}{2}x_2^2 f_{22}(0, 0). \end{aligned}$$

Differentiating:

$$\begin{aligned} r_1(x_1, x_2) &= f_1(0, 0) + x_1 f_{11}(0, 0) + x_2 f_{12}(0, 0) \\ r_2(x_1, x_2) &= f_2(0, 0) + x_1 f_{12}(0, 0) + x_2 f_{22}(0, 0) \\ r_{11}(x_1, x_2) &= f_{11}(0, 0) \\ r_{12}(x_1, x_2) &= f_{12}(0, 0) \\ r_{22}(x_1, x_2) &= f_{22}(0, 0) \end{aligned}$$

and so the image of r and its first and second partial derivatives at $(0, 0)$, agree with the image of f and its first and second partial derivatives at $(0, 0)$.

Thus, the second-degree Taylor approximation to $f(x_1, x_2)$ about $(0, 0)$ is

$$\begin{aligned} f(x_1, x_2) \simeq f(0, 0) & \\ &+ x_1 f_1(0, 0) + x_2 f_2(0, 0) \\ &+ \tfrac{1}{2}x_1^2 f_{11}(0, 0) + x_1 x_2 f_{12}(0, 0) + \tfrac{1}{2}x_2^2 f_{22}(0, 0). \end{aligned}$$

We have written this expression on three lines because it breaks up into three distinct parts. On the first line, we have a constant, on the second line, a linear functional on R^2, and on the third line, a quadratic form on R^2. Thus we can write

$$f(x_1, x_2) \simeq f(0, 0) + P(x_1, x_2) + Q(x_1, x_2) \qquad ((x_1, x_2) \in R^2)$$

where P is the linear functional specified by

$$P(x_1, x_2) = x_1 f_1(0, 0) + x_2 f_2(0, 0)$$

and Q is the quadratic form specified by

$$Q(x_1, x_2) = \tfrac{1}{2}x_1^2 f_{11}(0, 0) + x_1 x_2 f_{12}(0, 0) + \tfrac{1}{2}x_2^2 f_{22}(0, 0).$$

The linear functional P involves the first partial derivatives of f at $(0, 0)$, while the quadratic form Q involves the second partial derivatives at $(0, 0)$.

The polar form of Q is given by

$$\begin{aligned} G((x_1, x_2), (y_1, y_2)) &= \tfrac{1}{2}x_1 y_1 f_{11}(0, 0) + \tfrac{1}{2}x_1 y_2 f_{12}(0, 0) \\ &\quad + \tfrac{1}{2}x_2 y_1 f_{12}(0, 0) \\ &\quad + \tfrac{1}{2}x_2 y_2 f_{22}(0, 0) \\ &\qquad\qquad ((x_1, x_2), (y_1, y_2) \in R^2) \\ &= \tfrac{1}{2}(x_1 y_1 f_{11}(0, 0) + x_1 y_2 f_{12}(0, 0) \\ &\quad + x_2 y_1 f_{21}(0, 0) + x_2 y_2 f_{22}(0, 0)) \end{aligned}$$

since $f_{21} = f_{12}$.

We can write this

$$G((x_1, x_2), (y_1, y_2)) = \frac{1}{2} \sum_{i, j=1}^{2} x_i y_j f_{ij}(0, 0).$$

Exercise

Write down:

(i) the first-degree Taylor approximation about $(0, 0)$ to

$$\sin\left(\frac{\pi}{4} + x_1 + x_2\right) \qquad ((x_1, x_2) \in R^2);$$

(ii) the corresponding second-degree approximation.

Solution

The first and second partial derivatives of

$f(x_1, x_2) = \sin\left(\frac{\pi}{4} + x_1 + x_2\right)$ are:

$$f_1(x_1, x_2) = \cos\left(\frac{\pi}{4} + x_1 + x_2\right) \qquad f_2(x_1, x_2) = \cos\left(\frac{\pi}{4} + x_1 + x_2\right)$$

$$f_{11}(x_1, x_2) = -\sin\left(\frac{\pi}{4} + x_1 + x_2\right) \quad f_{12}(x_1, x_2) = -\sin\left(\frac{\pi}{4} + x_1 + x_2\right)$$

$$f_{22}(x_1, x_2) = -\sin\left(\frac{\pi}{4} + x_1 + x_2\right).$$

Thus:

(i) $$\sin\left(\frac{\pi}{4}+x_1+x_2\right) \simeq f(0,0)+x_1f_1(0,0)+x_2f_2(0,0)$$

$$= \sin\frac{\pi}{4}+x_1\cos\frac{\pi}{4}+x_2\cos\frac{\pi}{4}$$

$$= \frac{1}{\sqrt{2}}(1+x_1+x_2)$$

(ii) $$\sin\left(\frac{\pi}{4}+x_1+x_2\right) \simeq f(0,0)+x_1f_1(0,0)+x_2f_2(0,0)$$

$$+\frac{1}{2}x_1^2f_{11}(0,0)+x_1x_2f_{12}(0,0)$$

$$+\frac{1}{2}x_2^2f_{22}(0,0)$$

$$= \sin\frac{\pi}{4}+x_1\cos\frac{\pi}{4}+x_2\cos\frac{\pi}{4}$$

$$-\frac{1}{2}x_1^2\sin\frac{\pi}{4}-x_1x_2\sin\frac{\pi}{4}-\frac{1}{2}x_2^2\sin\frac{\pi}{4}$$

$$= \frac{1}{\sqrt{2}}\left(1+x_1+x_2-\frac{1}{2}x_1^2-x_1x_2-\frac{1}{2}x_2^2\right).$$

14.2.4 Summary of Section 14.2

In this section we defined the terms

quadratic form (page N160) * * *
polar form (page N161) *

Theorem

(9.1, page N161)
Every symmetric bilinear form f_s determines a unique quadratic form by the rule $q(\alpha) = f_s(\alpha, \alpha)$, and if $1+1 \neq 0$, every quadratic form determines a unique symmetric bilinear form $f_s(\alpha, \beta) = \frac{1}{2}[q(\alpha+\beta)-q(\alpha)-q(\beta)]$ from which it is in turn determined by the given rule. There is a one-to-one correspondence between symmetric bilinear forms and quadratic forms. * * *

Technique

Given a particular function of two variables, $f(x_1, x_2)$, find the quadratic Taylor approximation using the formula: * *

$$f(x_1,x_2) \simeq f(0,0)+\sum_{i=1}^{2}x_if_i(0,0)+\frac{1}{2}\sum_{i,j=1}^{2}x_ix_jf_{ij}(0,0)$$

Notation

(x) (page N161)

14.3 THE NORMAL FORM

14.3.0 Introduction

We saw in sub-section 14.1.2 that if two square matrices A and A' represent the same bilinear form with respect to different bases, then $A' = P^T AP$ where P is a matrix of transition. The matrices A and A' related in this way are said to be congruent and we verified in Exercise 2 of sub-section 14.1.2 that congruence is an equivalence relation. This means that once again it is useful to define a normal form for a matrix, namely the "simplest" matrix in the equivalence class.* As in the case of Hermite normal form, we shall see that the way to compute the normal form is to work in stages, making the matrix progressively simpler.

This particular type of normal form is, of course, only useful for a matrix that represents a bilinear form: the normal form appropriate to a particular discussion depends essentially on the effect of changes of bases. The Hermite normal form arose in the context of the representation of a linear transformation by a matrix, where we considered change of bases in the codomain only. The Jordan normal form arose in the same context, but we were considering linear transformations of a space to itself. Here the context is different: we have found that we can represent a bilinear form (with respect to some basis) by a matrix and the change of basis has an effect on the matrix which is different from the previous two.

To illustrate the usefulness of normal forms under the congruence relation, we will demonstrate what happens in the case of matrices representing symmetric bilinear forms on R^2.

Apart from the zero normal form, the normal forms that arise are

(i) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, (ii) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, (iii) $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$,

(iv) $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$.

If for each bilinear form q we consider the graph of the equation $q(x_1, x_2) = 1$, then the first of the above normal forms corresponds to a circle or ellipse, the second to a pair of straight lines, the third to a hyperbola, and there is no graph corresponding to the fourth normal form; in this case the solution set of $q(x_1, x_2) = 1$ is empty.

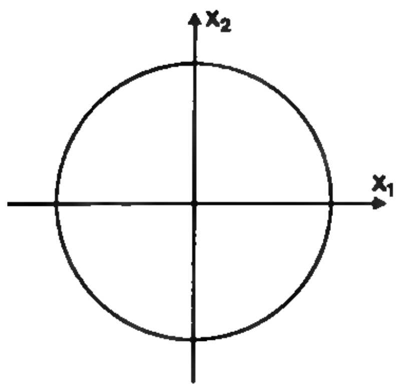
Exercise

Suppose that the above matrices represent bilinear forms with respect to the standard basis in R^2, and that a variable point in R^2 is denoted by (x_1, x_2). Write out the equation $q(x_1, x_2) = 1$ in terms of x_1, x_2 in each case, and draw its graph.

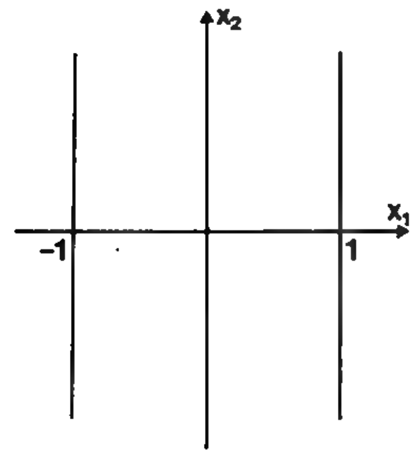
* You may find it of interest to read Section II-9, pages N74–78, which discusses the concept of normal form in general; we now have a few examples. We also discussed the general concept in *Unit 10, Jordan Normal Form.*

Solution

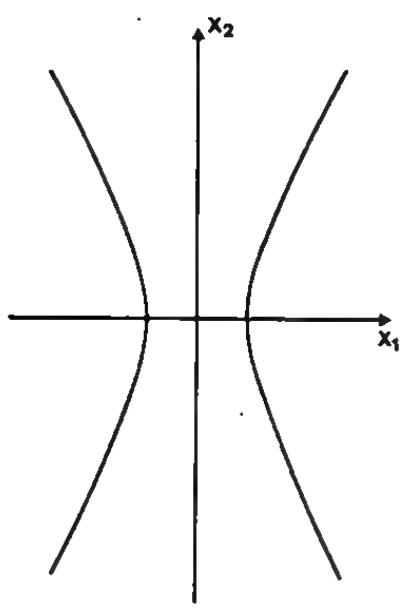
(i) $x_1^2 + x_2^2 = 1$.



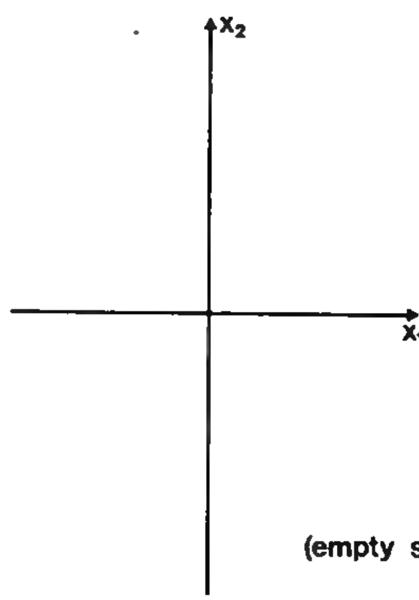
(ii) $x_1^2 = 1$, i.e. $x_1 = \pm 1$.



(iii) $x_1^2 - x_2^2 = 1$.



(iv) $-x_1^2 - x_2^2 = 1$.



(empty set)

14.3.1 Getting the Matrix Diagonal

The first step which we take in reducing a symmetric matrix to its normal form is to diagonalize it. (We consider symmetric matrices only, since among the many bilinear forms corresponding to a quadratic form, a symmetric bilinear form can always be chosen.)

Section IV-10 of N contains a theorem showing that any bilinear form can be diagonalized and an algorithm for doing it. We do not expect you to use this method, but the theorem is the subject of the next reading passage.

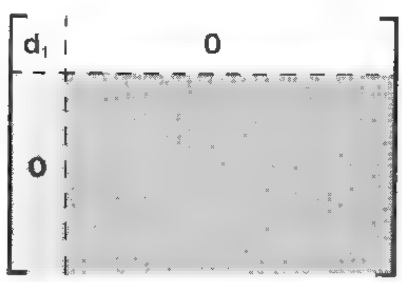
READ Section IV-10 starting on page N162 *as far as line* −5 (*the end of the proof*) *on page* N163.

Notes

(i) *line 10, page* N163 In less technical language, the proof shows that if $n > 1$, we can reduce the dimension of the problem by 1. Since we can repeat this technique until we get down to a 1-dimensional problem, and since a 1×1 matrix is automatically diagonal, we know that we can in principle tackle an n-dimensional problem for *any* finite n.
(ii) *Equation (10.1), page* N163 It is here that we use the fact that the matrix is symmetric.
(iii) *line 18: the sentence after Equation (10.1), page* N163 This is really a "proof by contradiction" in miniature. If we assume $q(\alpha) = 0$ for all α, then the matrix representing it with respect to any basis is 0. But we assumed a few lines back that $B \neq 0$. So there must exist an α such that $q(\alpha) \neq 0$.
(iv) *line 21, page* N163 The linear functional is $\alpha \longmapsto f_s(\alpha_1', \alpha)$.
(v) *line 23, page* N163 Since the linear functional is not the zero function, its rank is 1. The dimension of its domain is n, and therefore the dimension of its kernel W_1 is $n - 1$ (by ***Theorem 1.6*** on page N31).
(vi) *line 24, page* N163 "f_s restricted to W_1" means the bilinear form

$$(\alpha, \beta) \longmapsto f_s(\alpha, \beta) \qquad (\alpha \in W_1, \beta \in W_1)$$

with domain $W_1 \times W_1$ instead of $V \times V$.
(vii) *line 25, page* N163 "by assumption" that is, by our inductive hypothesis that the theorem has already been established for bilinear forms on spaces (here W_1) of dimension $n - 1$.
(viii) *line 26, page* N163 The notation "$2 \le i, j \le n$" means "$2 \le i \le n$ and $2 \le j \le n$"
(ix) *line 27, page* N163 In other words, the part of the matrix below shown unshaded contains only zeros (except for d_1) because we chose $\alpha_2', \ldots, \alpha_n'$ in the set $W_1 = \{\alpha : f_s(\alpha_1', \alpha) = 0\}$. The small square contains only the one entry d_1, and the large square is the matrix of the bilinear form "f_s restricted to W_1", and is therefore diagonal by our inductive hypothesis.



We next consider two methods of obtaining a diagonal form of a symmetric matrix, B:

(a) the method of elementary row and column operations,
(b) the method of completing the square.

Elementary Row and Column Operations

If P is the matrix of transition, the diagonal form B' is given by

$$B' = P^T B P.$$

Since P is non-singular, it is the product of elementary matrices. We can therefore get from B to BP by applying the corresponding column operations. Similarly, P^T is the product of elementary matrices, and we can get from BP to $P^T BP$ by applying the corresponding row operations.

Now it turns out that the required row operations correspond exactly to the required column operations, and must be performed in the same order. To see this, suppose that P is written out as a product of elementary matrices:

$$P = E_1 E_2 \dots E_n.$$

Then to find P^T, we must multiply the transposes of the elementary matrices together in the opposite order:

$$P^T = E_n^T \dots E_2^T E_1^T$$

Thus:

$$B' = E_n^T \dots E_2^T E_1^T B E_1 E_2 \dots E_n.$$

Now it is not difficult to see that if multiplying by E_1 on the right corresponds to multiplying column i by c, then $E_1^T = E_1$, and multiplying by E_1^T on the left corresponds to multiplying row i by c. Similarly, if multiplying by E_2 on the right corresponds to interchanging columns i and j, then $E_2^T = E_2$, and multiplying by E_2^T on the left corresponds to interchanging rows i and j. Finally, if multiplying by E_3 on the right corresponds to adding k times column i to column j, then (although $E_3^T \neq E_3$) multiplying by E_3^T on the left corresponds to adding k times row i to row j.

To get from B to

$$B' = E_n^T \dots E_2^T E_1^T B E_1 E_2 \dots E_n,$$

therefore, we must apply column and row operations in pairs. Applying a column operation and the corresponding row operation takes the matrix B into the matrix $E_1^T B E_1$; the next pair takes this into the matrix $E_2^T E_1^T B E_1 E_2$; and so on.

The next question to be settled is: how do we decide what column and row operations to perform? Perhaps the best way to answer this question is to go through an example.

Example

Find a diagonal form for $B = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

Step 1

Use each non-zero diagonal element to reduce to zero each element in the row and column occupied by that non-zero element. Starting with the first non-zero diagonal element, namely the top left-hand entry, $b_{11} = 1$, we reduce b_{12} and b_{21} to zero by subtracting column 1 from column 2, then subtracting row 1 from row 2:

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} B \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Step 2

The second and third diagonal elements are now zero, so we have to create a non-zero diagonal element in one or both of these positions. If $b_{ij} \neq 0$ $(i \neq j)$, add column j to column i, and row j to row i; this gives

$2b_{ij}$ in the ith place of the diagonal. In this case, then, we add column 3 to column 2, and row 3 to row 2:

$$\begin{bmatrix}1&0&0\\0&1&1\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\-1&1&0\\0&0&1\end{bmatrix}B\begin{bmatrix}1&-1&0\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\0&1&1\end{bmatrix}$$

$$=\begin{bmatrix}1&0&0\\-1&1&1\\0&0&1\end{bmatrix}B\begin{bmatrix}1&-1&0\\0&1&0\\0&1&1\end{bmatrix}=\begin{bmatrix}1&0&0\\0&2&1\\0&1&0\end{bmatrix}$$

Step 3

We can now use the 2 in the (22) position, to reduce to (23) and (32) elements to zero. We do this by subtracting $\frac{1}{2}$ column 2 from column 3, and $\frac{1}{2}$ row 2 from row 3:

$$\begin{bmatrix}1&0&0\\0&1&0\\0&-\frac{1}{2}&1\end{bmatrix}\begin{bmatrix}1&0&0\\-1&1&1\\0&0&1\end{bmatrix}B\begin{bmatrix}1&-1&0\\0&1&0\\0&1&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&-\frac{1}{2}\\0&0&1\end{bmatrix}$$

$$=\begin{bmatrix}1&0&0\\-1&1&1\\\frac{1}{2}&-\frac{1}{2}&\frac{1}{2}\end{bmatrix}B\begin{bmatrix}1&-1&\frac{1}{2}\\0&1&-\frac{1}{2}\\0&1&\frac{1}{2}\end{bmatrix}=\begin{bmatrix}1&0&0\\0&2&0\\0&0&-\frac{1}{2}\end{bmatrix}$$

Thus a diagonal form of B is $B'=\begin{bmatrix}1&0&0\\0&2&0\\0&0&-\frac{1}{2}\end{bmatrix}$, with transition matrix $P=\begin{bmatrix}1&-1&\frac{1}{2}\\0&1&-\frac{1}{2}\\0&1&\frac{1}{2}\end{bmatrix}$. Since this is the transition matrix, the underlying symmetric bilinear form has matrix representation B' with respect to the basis

$$\{(1,0,0),(-1,1,1),(\tfrac{1}{2},-\tfrac{1}{2},\tfrac{1}{2})\}$$

the coordinates of which are expressed in terms of the original basis.

In the above example we have explicitly recorded all the elementary matrices, although at each stage one is the transpose of the other. This acts as a check, but can of course be dropped.

Completing the Square

For simple quadratic forms the method of completing the square is easy and useful. The idea is to express the quadratic form as a linear combination of squares, and the method is to keep subtracting (or adding) appropriately constructed multiples of squares so as to leave one less variable in the quadratic form each time.

Example

Suppose we have the quadratic form

$$q((x))=x_1^2+2x_1x_2+2x_1x_3-x_2^2+3x_3^2$$

If we select a basis $\{\alpha_1,\alpha_2,\alpha_3\}$ of R^3 such that

$$q((x))=k_1x_1'^2+k_2x_2'^2+k_3x_3'^2$$

where

$$(x)=x_1'\alpha_1+x_2'\alpha_2+x_3'\alpha_3\qquad((x)\in R^3),$$

then the choice of $\{\alpha_1,\alpha_2,\alpha_3\}$ as a basis for R^3 has diagonalized the expression for $q((x))$. In practice, rather than look for $\alpha_1,\alpha_2,\alpha_3$ explicitly,

it is easier to look for expressions

$$\begin{aligned}
x_1' &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\
x_2' &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\
x_3' &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3
\end{aligned}$$

such that

$$x_1^2 + 2x_1x_2 + 2x_1x_3 - x_2^2 + 3x_3^2 = k_1x_1'^2 + k_2x_2'^2 + k_3x_3'^2$$

First, can we choose x_1' and k_1 so that

$$q((x)) - k_1x_1'^2$$

depends only on x_2 and x_3? In other words, can we choose $a_{11}x_1 + a_{12}x_2 + a_{13}x_3$ such that the coefficients of x_1^2, x_1x_2 and x_1x_3 in $(a_{11}x_1 + a_{12}x_2 + a_{13}x_3)^2$ are *proportional* to those in $q((x))$? If so, then all we do is subtract $x_1'^2$ multiplied by the relevant constant k_1 from $q((x))$ and we are left with an expression in x_2, x_3.

Now the terms involving x_1 in the expansion of $(a_{11}x_1 + a_{12}x_2 + a_{13}x_3)^2$ are:

$$\begin{aligned}
a_{11}^2x_1^2 + 2a_{11}a_{12}x_1x_2 + 2a_{11}a_{13}x_1x_3& \\
&= a_{11}(a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3).
\end{aligned}$$

Thus (provided $a_{11} \neq 0$) we can choose $k_1 = a_{11}$ and then get the right answer by choosing

$$\begin{aligned}
a_{11} &= \text{coefficient of } x_1^2 \text{ in } q((x)), \\
a_{12} &= \tfrac{1}{2} \text{ coefficient of } x_1x_2 \text{ in } q((x)), \\
a_{13} &= \tfrac{1}{2} \text{ coefficient of } x_1x_3 \text{ in } q((x)).
\end{aligned}$$

(Note the similarity to the rule of thumb we mentioned in sub-section 14.2.2 for finding the polar form of a quadratic form.)

In this particular case, then, we take

$$x_1' = x_1 + x_2 + x_3$$

and hence $k_1 = 1$.

This gives

$$\begin{aligned}
q((x)) - x_1'^2 &= q((x)) - (x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + 2x_2x_3 + 2x_1x_3) \\
&= -2x_2^2 + 2x_3^2 - 2x_2x_3.
\end{aligned}$$

To eliminate x_2 from the remaining expression on the right, we apply the same technique; $x_2' = a_{22}x_2 + a_{23}x_3$, where a_{22} is the coefficient of x_2^2, and a_{23} is half the coefficient of x_2x_3.

Thus

$$x_2' = -2x_2 - x_3$$

and hence $k_2 = -\frac{1}{2}$.

This gives

$$\begin{aligned}
q((x)) - x_1'^2 + \tfrac{1}{2}x_2'^2 &= q((x)) - x_1'^2 + \tfrac{1}{2}(4x_2^2 + 4x_2x_3 + x_3^2) \\
&= \tfrac{5}{2}x_3^2
\end{aligned}$$

Finally, therefore, we get

$$\begin{aligned}
q((x)) &= x_1'^2 - \tfrac{1}{2}x_2'^2 + \tfrac{5}{2}x_3'^2 \\
&= (x_1 + x_2 + x_3) - \tfrac{1}{2}(2x_2 + x_3)^2 + \tfrac{5}{2}x_3^2
\end{aligned}$$

(We have suppressed the minuses in x_2', since $(-1)^2 = 1$.)

The matrix of q with respect to the new basis (which we have not calculated explicitly) is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & \frac{5}{2} \end{bmatrix}.$$

This method of diagonalizing a quadratic form always works provided that, at each step in the process, there is an x^2 term left in the expression. If this fails to happen, then we have to alter our procedure and take an intermediate step to create non-zero elements on the diagonal.

Example

$$q((x)) = x_1x_2 + x_2x_3 + x_3x_1$$

In this case, we cannot straightforwardly eliminate x_1 by the subtraction or addition of a single square, since to eliminate the x_1x_2 and x_1x_3 terms we would have to create an x_1^2 term. The technique here is to define the intermediate variables

$$\begin{aligned} x_1' &= x_1 + x_2 \\ x_2' &= x_1 - x_2 . \end{aligned}$$

We clearly still have linear independence of the linear functionals x_1', x_2' and x_3, and the expression for $q((x))$ becomes

$$\begin{aligned} q((x)) &= (\tfrac{1}{2}x_1' + \tfrac{1}{2}x_2')(\tfrac{1}{2}x_1' - \tfrac{1}{2}x_2') + (\tfrac{1}{2}x_1' - \tfrac{1}{2}x_2')x_3 \\ &\qquad + x_3(\tfrac{1}{2}x_1' + \tfrac{1}{2}x_2') \\ &= \tfrac{1}{4}x_1'^2 - \tfrac{1}{4}x_2'^2 + x_1'x_3. \end{aligned}$$

Now we proceed exactly as before. Define

$$x_1'' = \tfrac{1}{4}x_1' + \tfrac{1}{2}x_3$$

and we have diagonalized q:

$$\begin{aligned} q((x)) &= 4x_1''^2 - \tfrac{1}{4}x_2'^2 - 3x_3^2 \\ &= 4(\tfrac{1}{4}x_1 + \tfrac{1}{4}x_2 + \tfrac{1}{2}x_3)^2 - \tfrac{1}{4}(x_1 - x_2)^2 - x_3^2 \\ &= \tfrac{1}{4}(x_1 + x_2 + 2x_3)^2 - \tfrac{1}{4}(x_1 - x_2)^2 - x_3^2 . \end{aligned}$$

READ from line −14 on page N166 *to the end of the section on page* N167.

Note

line −2, page N166 If $X^TBX = \sum x_i b_{ij} x_j$ and B is symmetric, then $b_{ij} = b_{ji} = \frac{1}{2}$ (the coefficient of the x_ix_j term in the quadratic expression).

For example

$$\begin{aligned} x^2 + 3xy + y^2 &= x^2 + \tfrac{3}{2}xy + \tfrac{3}{2}yx + y^2 \\ &= [x\ y]\begin{bmatrix} 1 & \frac{3}{2} \\ \frac{3}{2} & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} \end{aligned}$$

Exercises

1. Use elementary row and column operations to find diagonal forms for the following symmetric matrices. (They are the matrices representing the quadratic forms in the examples worked by the method of completing the square.)

   (i) $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$ (ii) $B = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$

2. Reduce to diagonal form by the method of completing the square the quadratic forms given in Exercise 1, page N162. (Do as many as you feel you need to.)

Solutions

1. (i)

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Subtract row 1 from row 2, and then subtract column 1 from column 2:

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & -2 & -1 \\ 1 & -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Subtract row 1 from row 3, and then subtract column 1 from column 3:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Subtract $\frac{1}{2}$row 2 from row 3, and then subtract $\frac{1}{2}$column 2 from column 3:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & \frac{5}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix} A \begin{bmatrix} 1 & -1 & -\frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}$$

(N.B. The diagonal form obtained by completing the square can be obtained by: row 2 becomes $\frac{1}{2}$row 2, and then column 2 becomes $\frac{1}{2}$column 2.)

(ii)

$$\begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} B \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Since there is no non-zero element on the diagonal, add row 2 to row 1, and then add column 2 to column 1:

$$\begin{bmatrix} 1 & \frac{1}{2} & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & \frac{1}{2} & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} B \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Subtract $\frac{1}{2}$row 1 from row 2, and then subtract $\frac{1}{2}$column 1 from column 2:

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & -\frac{1}{4} & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} B \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Subtract row 1 from row 3, and subtract column 1 from column 3:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \\ 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ -1 & -1 & 1 \end{bmatrix} B \begin{bmatrix} 1 & -\frac{1}{2} & -1 \\ 1 & \frac{1}{2} & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

2. (a) $2x^2 + 3xy + 6y^2$: subtract $\frac{1}{2}(2x + \frac{3}{2}y)^2$.

This leaves $4\frac{7}{8}y^2$. Thus

$$2x^2 + 3xy + 6y^2 = \tfrac{1}{2}(2x + \tfrac{3}{2}y)^2 + 4\tfrac{7}{8}y^2.$$

(b) $8xy + 4y^2$: subtract $\frac{1}{4}(4x + 4y)^2 = 4(x + y)^2$:

This leaves $-4x^2$. Thus

$$8xy + 4y^2 = 4(x + y)^2 - 4x^2.$$

(c) $x^2 + 2xy + 4xz + 3y^2 + yz + 7z^2$: subtract $(x + y + 2z)^2$.

This leaves $2y^2 - 3yz + 3z^2$: subtract $\frac{1}{2}(2y - \frac{3}{2}z)^2$. This leaves $1\frac{7}{8}z^2$. Thus

$$\begin{aligned} &x^2 + 2xy + 4xz + 3y^2 + yz + 7z^2 \\ &\qquad = (x + y + 2z)^2 + \tfrac{1}{2}(2y - \tfrac{3}{2}z)^2 + 1\tfrac{7}{8}z^2. \end{aligned}$$

(d) $4xy = (x + y)^2 - (x - y)^2$

(e) $x^2 + 4xy + 4y^2 + 2xz + z^2 + 4yz$:

subtract $(x + 2y + z)^2$. This leaves 0: thus

$$\begin{aligned} &x^2 + 4xy + 4y^2 + 2xz + z^2 + 4yz \\ &\qquad = (x + 2y + z)^2 \end{aligned}$$

(f) $x^2 + 4xy - 2y^2 = (x + 2y)^2 - 6y^2$.

(g) $x^2 + 6xy - 2y^2 - 2yz + z^2$: subtract $(x + 3y)^2$.

This leaves $-11y^2 - 2yz + z^2$: subtract $(y - z)^2$. This leaves $-12y^2$. Thus

$$\begin{aligned} &x^2 + 6xy - 2y^2 - 2yz + z^2 \\ &\qquad = (x + 3y)^2 + (y - z)^2 - 12y^2. \end{aligned}$$

14.3.2 Getting a Diagonal Matrix into Normal Form

Having diagonalized the matrix of a bilinear form over V, we want to simplify the diagonal elements.

The extent of the simplification depends on whether V is over the field of real numbers or the field of complex numbers. (We will not consider other fields.) Having got the matrix into diagonal form, we want to keep it there; this means that we can only multiply our basis elements by non-zero scalars. How much freedom of action does this give us? Suppose $\{\alpha_1, \ldots, \alpha_n\}$ is a basis of V with respect to which the matrix $B = [b_{ij}]$ of a bilinear form f is diagonal. Then if we change the basis element α_1 to $a\alpha_1$, where a is a non-zero scalar, we find that b_{11} is replaced by a^2b_{11}, for $f(a\alpha_1, a\alpha_1) = a^2f(\alpha_1) = a^2b_{11}$. Similarly, we can multiply any other diagonal element b_{ii} by a non-zero scalar. We still have the off-diagonal entries $f(a\alpha_i, \alpha_j)$ equal to zero.

Suppose first that our field of scalars is R. Then the factor a^2 multiplying a diagonal entry in this way is positive; furthermore, any positive number has a square root. Thus any non-zero element b_{ii} on the diagonal can be reduced to either 1 or -1 depending on the sign of b_{ii}, by multiplying the corresponding basis element by

$$\frac{1}{\sqrt{b_{ii}}} \quad \left(\text{or } \frac{1}{\sqrt{-b_{ii}}} \text{ if } b_{ii} \text{ is negative}\right).$$

On the other hand, zero elements off the diagonal remain zero. Finally, we can re-arrange the basis elements so that the $+1$s come first, followed by the -1s, followed by the 0s. This gives the following result.

Theorem

If V is a vector space over R, the normal form for the matrix corresponding to a symmetric bilinear form f is a diagonal matrix whose diagonal entries consist entirely of 1s, -1s and 0s, which may for convenience be arranged in that order.

If the field is C, however, the story is somewhat different; *every* element of C has a square root. Thus *every* non-zero diagonal entry b_{ii} in the matrix can be brought to $+1$ by multiplying the corresponding basis element by $\dfrac{1}{\sqrt{b_{ii}}}$. We therefore have the following result.

Theorem

If V is a vector space over C, the normal form for the matrix corresponding to a symmetric bilinear form f is a diagonal matrix whose diagonal entries consist entirely of 1s and 0s, with the 1s preceding the 0s for convenience.

Exercises

1. Write down the normal forms of the matrices of the quadratic forms in the final exercise of sub-section 14.3.1. Their diagonal forms are given in the solution to that exercise. *Assume that the underlying field is R.*
2. Repeat Exercise 1 but take the underlying field to be C.

Solutions

1. (a) $\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$ (b) $\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$ (c) $\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$

   (d) $\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$ (e) $\begin{bmatrix}1 & 0 & 0\\0 & 0 & 0\\0 & 0 & 0\end{bmatrix}$ (f) $\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$

   (g) $\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & -1\end{bmatrix}$

2. As above but with -1 replaced by 1 wherever it occurs, e.g.

   (b) $\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$.

14.3.3 Uniqueness of the Normal Form

In sub-section 14.3.2 we showed that, from a given diagonal form, we can get to a unique normal form (different for the fields R and C). However, we did not show that different diagonal forms for the same quadratic form will lead to the same normal form. What we have to show is that any diagonal form contains a unique number of positive, negative and zero diagonal elements (in the case of R), or a unique number of non-zero and zero elements (in the case of C). If we can show this, then the normal form will be unique.

First of all, it is easy enough to establish that the number of zeros on the diagonal in the diagonal form of the matrix is unique. Since the diagonal form of B is P^TBP for some non-singular matrix P, and since P^T is also non-singular, it follows (Theorem 3.7 on page N48) that the rank of B is equal to the rank of P^TBP; and this is, of course, equal to the number of

non-zero diagonal elements of P^TBP. This completely solves the uniqueness problem in the case of the complex field, C, as all we want to know is the number of non-zero, and the number of zero, diagonal elements in the diagonal form.

Definition

The *rank* of a quadratic form (or of the corresponding polar form) is equal to the rank of the matrix representing it.

For the case of the real field, we want a further result; of the non-zero diagonal elements, we want to know that the number of *positive* elements is unique. This is proved in the first part of Section IV-11 of N.

READ from the beginning of Section IV-11 on page N168 *to line 5 on page* N169.

Notes

(i) ***Theorem 11.1, page*** N168 The proof depends on a theorem:

$$\dim U + \dim W = \dim (U + W) + \dim (U \cap W),$$

which we have not covered in this course. It is ***Theorem 4.8*** on page N22. Here $U + W$ means the subspace $\{\alpha + \beta : \alpha \in U \text{ and } \beta \in W\}$.
(ii) *line* -4, *page* N168 r is the rank of q.
(iii) *lines* -3 *to* -1, *page* N168 It is important to remember what non-negative semi-definite and positive definite forms are, less important to remember what a signature is. A non-negative semi-definite quadratic form has no -1s (but possibly some 0s) in its normal form; a positive definite quadratic form has all 1s in its normal form. In matrix language, A is non-negative semi-definite if and only if $X^TAX \geq 0$ for all one-column matrices X, and is positive definite if and only if, *in addition*, $X^TAX = 0$ implies

$$X = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Exercises

1. Exercise 1, page N170.
2. Exercise 2, page N170.
3. Exercise 3, page N170.
4. Exercise 4, page N170.

(Use the results of the exercises of the preceding sub-section of this text to help you with Exercise 1. In Exercise 2, consider the cases $a \neq 0$ and $a = 0$ separately.)

Solutions

1.

	Rank	*Signature*
(a)	2	2
(b)	2	0
(c)	3	3
(d)	2	0
(e)	1	1
(f)	2	0
(g)	3	1

2. There are two cases to consider: $a \neq 0$ and $a = 0$.

Case (i): $a \neq 0$. Completing the square gives

$$Q(x, y) = \frac{1}{a}\left(ax + \frac{b}{2}y\right)^2 + \left(c - \frac{b^2}{4a}\right)y^2$$

which is positive definite if and only if the coefficients

$$\frac{1}{a} \quad \text{and} \quad c - \frac{b^2}{4a}$$

are positive, i.e. if and only if $a > 0$ and $b^2 - 4ac < 0$.

Case (ii): $a = 0$

$$Q(x, y) = bxy + cy^2$$

We have to show that $Q(x, y)$ can *never* be positive definite in this case. If c is also zero, then completing the square gives

$$Q(x, y) = bxy = \frac{b}{4}[(x+y)^2 - (x-y)^2].$$

which cannot be positive definite. If $c \neq 0$, then

$$Q(x, y) = \frac{1}{c}\left(\frac{b}{2}x + cy\right)^2 - \frac{b^2}{4c}x^2$$

and either $\frac{1}{c} < 0$ or $\frac{-b^2}{4c} \leq 0$.

Thus if $a = 0$, $Q(x, y)$ can never be positive-definite.

3. The normal form of A is I; thus there is a real non-singular matrix Q such that $Q^T A Q = I$. If we left-multiply each side by $(Q^{-1})^T$ and right-multiply each side by Q^{-1}, we get

$$A = (Q^{-1})^T Q^{-1}.$$

Thus Q^{-1} is the required matrix.

4. If we pre-multiply $A^T A$ by the non-singular matrix $P^T = (A^{-1})^T$ and post-multiply it by the non-singular matrix $P = A^{-1}$, we get

$$P^T A^T A P = I.$$

Thus $A^T A$ has I as its normal form, and is therefore positive definite.

14.3.4 An Application of Real Quadratic Forms

Let us look again at the example in sub-section 14.2.3, (the two-variable Taylor expansion). We had

$$f(x_1, x_2) \simeq f(0, 0) + P(x_1, x_2) + Q(x_1, x_2) \tag{1}$$

where P is a linear functional on R^2, and Q is a quadratic form on R^2 (see page 25). Suppose $f_1(0, 0) = f_2(0, 0) = 0$, so that $P(0, 0) = 0$. Then, as we showed in *Unit M100 15*, f has a stationary value at $(0, 0)$, and may have a local maximum or minimum at $(0, 0)$. We did not discuss in the above unit, however, the technique of working out from the second partial derivatives what kind of stationary value it is. For functions of *one* variable, the technique is: if $f'(x_0) = 0$, then there is a stationary value at x_0, which is a local maximum if $f''(x_0) < 0$ and a local minimum if $f''(x_0) > 0$. If $f''(x_0) = 0$, this classification method breaks down and another method must be used.

An analogous situation exists for functions of two variables. For simplicity of notation, we will suppose the stationary value is at $x_1 = x_2 = 0$; then Equation (1) becomes

$$f(x_1, x_2) \simeq f(0, 0) + Q(x_1, x_2) \tag{2}$$

since

$$P(x_1, x_2) = x_1 f_1(0, 0) + x_2 f_2(0, 0).$$

Now let $\{\alpha_1, \alpha_2\}$ be a basis of R^2 with respect to which Q is in normal form; let

$$(x_1, x_2) = u_1\alpha_1 + u_2\alpha_2 \qquad ((x_1, x_2) \in R^2)$$

so that u_1 and u_2 are the new variables with respect to which we express Equation (2).

Then Equation (2) has one of the following forms.

If the normal form of Q is $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, then

$$f(x_1, x_2) \simeq f(0, 0) + u_1^2 + u_2^2. \tag{3}$$

If the normal form of Q is $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, then

$$f(x_1, x_2) \simeq f(0, 0) + u_1^2 - u_2^2. \tag{4}$$

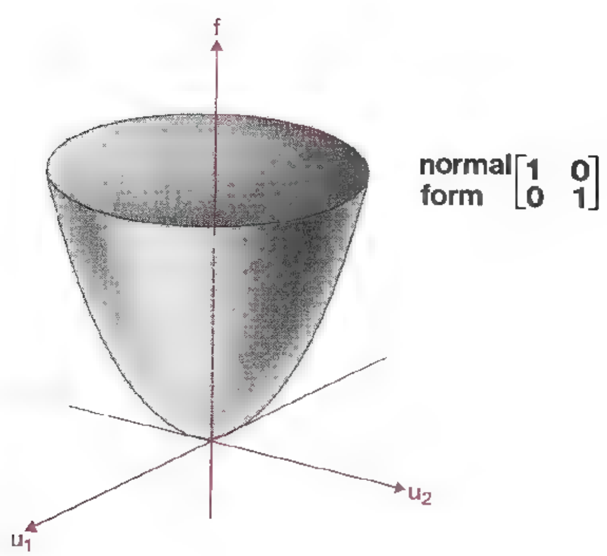
If the normal form of Q is $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$, then

$$f(x_1, x_2) \simeq f(0, 0) - u_1^2 - u_2^2. \tag{5}$$

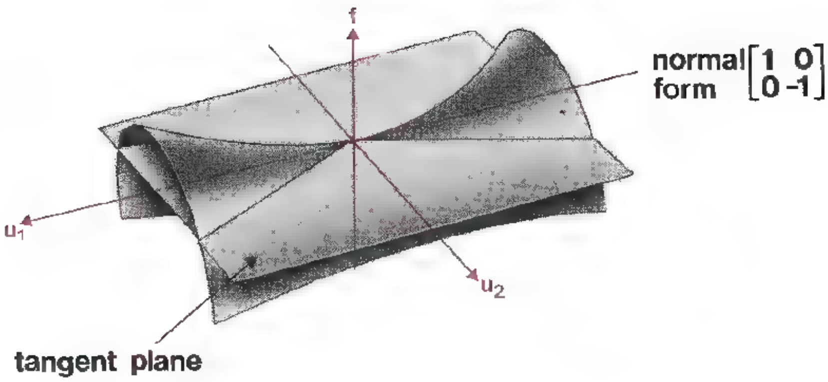
If the normal form of Q is $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, then

$$f(x_1, x_2) \simeq f(0, 0) + u_1^2. \tag{6}$$

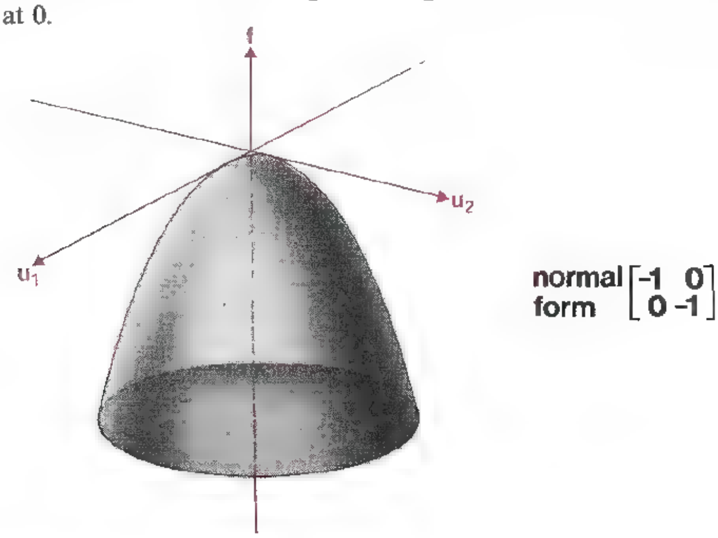
Thus, if the (real) normal form of Q is $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, i.e. if Q is positive definite, then there is a *local minimum* at (0, 0).



In the case of the normal form $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $f(x_1, x_2)$ increases with increasing u_1 if we go away from (0, 0) in the direction of α_1, but decreases with increasing u_2 if we go away from (0, 0) in the direction of α_2. Thus we get a *saddle point* at (0, 0) shown below.



In the case of the normal form $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$, we clearly get a *local maximum* at 0.



For Equation (6), we do not know whether $f(x_1, x_2)$ increases or decreases in the α_2 direction; the approximation is not good enough to tell us. The same goes for the normal forms: $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}$. Thus we can classify the stationary point using Q, whenever the rank of Q is 2, and we fail whenever the rank is less than 2.

Example

Classify the stationary point at (0, 0) of the function

$$f: (x_1, x_2) \longmapsto 4 + x_1^2 - 3x_1x_2 + x_2^2 + x_1^3 + x_1^2 x_2 \qquad ((x_1, x_2) \in R^2)$$

The quadratic approximation is (dropping the terms of higher than quadratic degree)

$$f(x_1, x_2) \simeq 4 + 0 + Q(x_1, x_2)$$

with

$$\begin{aligned} Q(x_1, x_2) &= x_1^2 - 3x_1x_2 + x_2^2 \\ &= (x_1 - \tfrac{3}{2}x_2)^2 - \tfrac{5}{4}x_2^2 \end{aligned}$$

(completing the square).

The normal form is therefore $\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$ and consequently the stationary point is a saddle point.

Example

Classify the stationary point at (0, 0) of the function

$$f: (x_1, x_2) \longmapsto \cos(x_1 + x_2) \qquad ((x_1, x_2) \in R^2)$$

The first stage is to calculate the first and second partial derivatives:

$$\begin{aligned}
f_1(x_1, x_2) &= -\sin(x_1 + x_2)\\
f_2(x_1, x_2) &= -\sin(x_1 + x_2)\\
f_{11}(x_1, x_2) &= -\cos(x_1 + x_2)\\
f_{12}(x_1, x_2) &= -\cos(x_1 + x_2)\\
f_{22}(x_1, x_2) &= -\cos(x_1 + x_2)
\end{aligned}$$

Thus

$$\begin{aligned}
f(x_1, x_2) \simeq f(0, 0) &+ x_1 f_1(0, 0) + x_2 f_2(0, 0)\\
&+ \tfrac{1}{2}x_1^2 f_{11}(0, 0) + x_1 x_2 f_{12}(0, 0) + \tfrac{1}{2}x_2^2 f_{22}(0, 0)\\
&= f(0, 0) + x_1 \sin(0) + x_2 \sin(0)\\
&\quad - \tfrac{1}{2}x_1^2 \cos(0) - x_1 x_2 \cos(0) - \tfrac{1}{2}x_2^2 \cos(0)\\
&= f(0, 0) - \tfrac{1}{2}x_1^2 - x_1 x_2 - \tfrac{1}{2}x_2^2\\
&= f(0, 0) - \tfrac{1}{2}(x_1 + x_2)^2
\end{aligned}$$

and so the quadratic form has $\begin{bmatrix}-1 & 0\\0 & 0\end{bmatrix}$ as a normal form, and this method does not enable us to classify the stationary point in this case.

The method is more successful in the following cases, which we give as an exercise.

Exercise

Classify the stationary point (0, 0) as a local maximum, local minimum, or saddle point for the following functions.

(i) $f: (x_1, x_2) \longmapsto \cos(x_1 + x_2) + \cos(x_1 - x_2) \qquad ((x_1, x_2) \in R^2)$

(ii) $g: (x_1, x_2) \longmapsto \cos(x_1 + x_2) - \cos(x_1 - x_2) \qquad ((x_1, x_2) \in R^2)$

Solution

(i)

$$\begin{aligned}
f_1(x_1, x_2) &= -\sin(x_1 + x_2) - \sin(x_1 - x_2)\\
f_2(x_1, x_2) &= -\sin(x_1 + x_2) + \sin(x_1 - x_2)\\
f_{11}(x_1, x_2) &= -\cos(x_1 + x_2) - \cos(x_1 - x_2)\\
f_{12}(x_1, x_2) &= -\cos(x_1 + x_2) + \cos(x_1 - x_2)\\
f_{22}(x_1, x_2) &= -\cos(x_1 + x_2) - \cos(x_1 - x_2)
\end{aligned}$$

Thus

$$\begin{aligned}
&f_1(0, 0) = f_2(0, 0) = 0,\\
&f_{11}(0, 0) = -2\\
&f_{12}(0, 0) = 0\\
&f_{22}(0, 0) = -2\\
&f(x_1, x_2) \simeq f(0, 0) - x_1^2 - x_2^2.
\end{aligned}$$

The normal form is $\begin{bmatrix}-1 & 0\\0 & -1\end{bmatrix}$ and hence the stationary point (0, 0) is a local maximum.

(ii) Here,

$$
\begin{aligned}
g_{11}(0,0) &= -\cos(0) + \cos(0) = 0\\
g_{12}(0,0) &= -\cos(0) - \cos(0) = -2\\
g_{22}(0,0) &= -\cos(0) + \cos(0) = 0\\
g(x_1, x_2) &\simeq g(0,0) - 2x_1x_2\\
&= g(0,0) + \tfrac{1}{2}(x_1 - x_2)^2 - \tfrac{1}{2}(x_1 + x_2)^2
\end{aligned}
$$

The normal form is $\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$ and hence the stationary point (0, 0) is a saddle point.

14.3.5 Summary of Section 14.3

In this section we defined the terms

rank	(page N164)	*
signature	(page N168)	*
non-negative semi-definite	(page N168)	* * *
positive definite	(page N168)	* * *

Theorems

1. (10.1, page N163)
For a given symmetric matrix B over a field F (in which $1 + 1 \neq 0$), there is a non-singular matrix P such that P^TBP is a diagonal matrix. In other words, if f_s is the underlying symmetric bilinear (polar) form, there is a basis $A' = \{\alpha_1', \ldots, \alpha_n'\}$ of V such that $f_s(\alpha_i', \alpha_j') = 0$ whenever $i \neq j$. * * *

2. (page C36)
If V is a vector space over R, the normal form for the matrix corresponding to a symmetric bilinear form f is the diagonal matrix: * * *

$$
\begin{bmatrix}
1 & & & & & & & \\
& \ddots & & & & & & \\
& & 1 & & & & & \\
& & & -1 & & & & \\
& & & & \ddots & & & \\
& & & & & -1 & & \\
& & & & & & 0 & \\
& & & & & & & \ddots \\
& & & & & & & & 0
\end{bmatrix}
$$

3. (page C36)
If V is a vector space over C, the normal form for the matrix corresponding to a symmetric bilinear form f is the diagonal matrix: *

$$
\begin{bmatrix}
1 & & & & \\
& \ddots & & & \\
& & 1 & & \\
& & & 0 & \\
& & & & \ddots \\
& & & & & 0
\end{bmatrix}.
$$

4. (Result from pages C36-7 and Theorem 11.1, page N168)
The number of positive and negative and null elements in a normal form matrix representing a quadratic form over R is unique. * *
The number of zero and non-zero elements in a normal form matrix representing a quadratic form over C is unique.

Techniques

1. Given a quadratic form over R or C, find its normal form by the method of row and column operations and by completing the square. * * *
2. Determine the basis for a diagonal form for a quadratic form. *
3. Given a particular $f(x_1, x_2)$, say what you can about maxima and minima at (0, 0). *

14.4 SUMMARY OF THE UNIT

As far as theory is concerned, this unit is a logical extension of *Unit 12, Linear Functionals and Duality*. The idea of a linear functional leads on to the idea of a bilinear functional, or bilinear form, and from there to the concept of a quadratic form, which is non-linear. There is a one-one correspondence between quadratic and symmetric bilinear forms which proves to be a useful tool, enabling us to analyse certain non-linear problems using linear techniques.

In the first section we looked at various kinds of bilinear forms and their matrix representation. We discovered that the matrices representing the same bilinear form with respect to different bases exhibit an equivalence relation called congruency.

The second section investigated the relationship between a quadratic form q and its corresponding symmetric bilinear form, called the polar form of q.

In the third section we developed two techniques for finding a simple representative matrix for a bilinear form, termed the normal form. Finally we applied the theory covered in this unit to analyse the stationary points of a suitably differentiable real-valued function of two variables.

Definitions

bilinear form	(page N156)	* * *
symmetric bilinear form	(page N158)	* * *
anti-symmetric bilinear form	(page C14)	*
skew-symmetric bilinear form	(page N158)	* * *
symmetric matrix	(page N158)	* * *
skew-symmetric matrix	(page N159)	* * *
congruent	(page N158)	* * *
quadratic form	(page N160)	* * *
polar form	(page N161)	*
rank	(page N164)	*
signature	(page N168)	*
non-negative semi-definite	(page N168)	* * *
positive definite	(page N168)	* * *

Theorems

1. (8.1, page N158)
A bilinear form f is symmetric if and only if any matrix B representing f has the property $B^T = B$. * * *

2. (8.2, page N158)
If a bilinear form f is skew-symmetric, then any matrix B representing f has the property $B^T = -B$. * * *

3. (8.3, page N158)
If $1 + 1 \neq 0$ and the matrix B representing f has the property $B^T = -B$, then f is skew-symmetric. * * *

4. (8.4, page N159)
If $1 + 1 \neq 0$, every bilinear form can be represented uniquely as a sum of a symmetric bilinear form and a skew-symmetric bilinear form. * * *

5. (9.1, page N161)
Every symmetric bilinear form f_s determines a unique quadratic form by the rule $q(\alpha) = f_s(\alpha, \alpha)$, and if $1 + 1 \neq 0$, every quadratic form determines a unique symmetric bilinear form $f_s(\alpha, \beta) = \frac{1}{2}[q(\alpha + \beta) - q(\alpha) - q(\beta)]$ from which it is in turn determined by the given rule. There is a one-to-one correspondence between symmetric bilinear forms and quadratic forms. * * *

6. (10.1, page N163)
For a given symmetric matrix B over a field F (in which $1 + 1 \neq 0$), there is a non-singular matrix P such that $P^T BP$ is a diagonal matrix. In other words, if f_s is the underlying symmetric bilinear (polar) form, there is a basis $A' = \{\alpha'_1, \ldots, \alpha'_n\}$ of V such that $f_s(\alpha'_i, \alpha'_j) = 0$ whenever $i \neq j$. * * *

7. (page C36)
If V is a vector space over R, the normal form for the matrix corresponding to a symmetric bilinear form f is the diagonal matrix: * * *

$$\begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{bmatrix}$$

8. (page C36)
If V is a vector space over C, the normal form for the matrix corresponding to a symmetric bilinear form f is the diagonal matrix: *

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

9. (Result from pages C36-7 and Theorem 11.1, page N168)
The number of positive and negative and null elements in a normal form matrix representing a quadratic form over R is unique. * *
The number of zero and non-zero elements in a normal form matrix representing a quadratic form over C is unique.

Techniques

1. Given a bilinear form f, find f_s, f_{ss}. * * *
2. Given a particular function of two variables, find the quadratic Taylor approximation. * *
3. Given a quadratic form over R or C find its normal form by the method of row and column operations and by completing the square. * * *
4. Determine the basis for a diagonal form of a quadratic form. *
5. Given a particular $f(x_1, x_2)$, say what you can about maxima and minima at (0, 0). *

Notation

$f_s(\alpha, \beta)$	(page N159)
$f_{ss}(\alpha, \beta)$	(page N159)
(x)	(page N161)

14.5 SELF-ASSESSMENT

Self-assessment Test

This Self-assessment Test is designed to help you test your understanding of the unit. It can also be used, together with the summary of the unit, for revision. The answers to these questions will be found on the next non-facing page. We suggest that you complete the whole test before looking at the answers.

1. Let f be the following bilinear form on R^3:

$$f((x), (y)) = x_1y_1 - x_1y_2 + x_2y_1 + x_2y_3 + 3x_3y_2 - x_3y_3 \qquad ((x), (y) \in R^3)$$

Calculate:

(i) the symmetric part, f_s, of f, and its matrix;
(ii) the skew-symmetric part, f_{ss}, of f, and its matrix;
(iii) the quadratic forms on R^3 corresponding to:

(a) f
(b) f_s
(c) f_{ss}

2. Show that the following formula gives the polar form for a quadratic form q on R^n:

$$f_s((x), (y)) = \frac{1}{2}\sum_{i=1}^{n} y_i q_i((x))$$

where $q_i((x))$ is the partial derivative of q with respect to x_i.

3. Show that if A is a matrix with real entries, then A^TA is the matrix of a real non-negative semi-definite quadratic form.

4. Determine whether the stationary point at (0, 0) of the following function is a local maximum, a local minimum, or a saddle point.

$$f:(x_1, x_2) \longmapsto e^{x_1+x_2} + e^{-x_1-x_2} + 2e^{x_1-x_2} + 2e^{-x_1+x_2} \qquad ((x_1, x_2) \in R^2)$$

Solutions to Self-assessment Test

1. (i) $f_s((x),(y)) = x_1y_1 + 2x_2y_3 + 2x_3y_2 - x_3y_3$;

$$\text{matrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 2 & -1 \end{bmatrix}.$$

(ii) $f_a((x),(y)) = -x_1y_2 + x_2y_1 - x_2y_3 + x_3y_2$;

$$\text{matrix} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

(iii) (a) $q((x)) = x_1^2 + 4x_2x_3 - x_3^2$
(b) $q((x)) = x_1^2 + 4x_2x_3 - x_3^2$
(c) $q((x)) = 0$

2. If a_i is the coefficient of the x_i^2 term in q, and b_{ij} the coefficient of the x_ix_j term ($i \neq j$), for each i, j, then the partial derivative with respect to x_i will have terms in $x_1, x_2, \ldots, x_n$. The coefficient of the x_i term will be $2a_i$, and the coefficient of the x_j term ($j \neq i$) will be b_{ij}. If we multiply this derivative by $\frac{1}{2}y_i$ and sum over i, we get a bilinear form whose x_iy_i term has coefficient a_i, and whose x_iy_j term ($i \neq j$) has coefficient $\frac{1}{2}b_{ij}$. This is the same as the formula derived in sub-section 14.2.2 for the polar form of q.

3. If A is $m \times n$, then A^TA is the product of an $n \times m$ with an $m \times n$ matrix, which is defined, and is an $n \times n$ matrix. It therefore defines a quadratic form on R^n, given by

$$q((x)) = X^T(A^TA)X$$

where X is the one-column matrix corresponding to (x). If we let $Y = (y_1, \ldots, y_m)$ be the one-column matrix AX, then

$$\begin{aligned} q((x)) &= Y^TY \\ &= \sum_{i=1}^{m} y_i^2 \\ &\geq 0 \qquad \text{for all } (x). \end{aligned}$$

Thus A^TA defines a non-negative semi-definite quadratic form.

4.
$$\begin{aligned} f_1(x_1, x_2) &= e^{x_1+x_2} - e^{-x_1-x_2} + 2e^{x_1-x_2} - 2e^{-x_1+x_2} \\ f_2(x_1, x_2) &= e^{x_1+x_2} - e^{-x_1-x_2} - 2e^{x_1-x_2} + 2e^{-x_1+x_2} \\ f_{11}(x_1, x_2) &= e^{x_1+x_2} + e^{-x_1-x_2} + 2e^{x_1-x_2} + 2e^{-x_1+x_2} \\ f_{12}(x_1, x_2) &= e^{x_1+x_2} + e^{-x_1-x_2} - 2e^{x_1-x_2} - 2e^{-x_1+x_2} \\ f_{22}(x_1, x_2) &= e^{x_1+x_2} + e^{-x_1-x_2} + 2e^{x_1-x_2} + 2e^{-x_1+x_2} \end{aligned}$$

Thus:

$$\begin{aligned} f_1(0,0) &= 0 \\ f_2(0,0) &= 0 \\ f_{11}(0,0) &= 6 \\ f_{12}(0,0) &= -2 \\ f_{22}(0,0) &= 6 \end{aligned}$$

The quadratic Taylor approximation to f about (0, 0) is therefore

$$f(x_1, x_2) \simeq f(0,0) + x_1 f_1(0,0) + x_2 f_2(0,0) + \tfrac{1}{2}x_1^2 f_{11}(0,0)$$
$$+ x_1 x_2 f_{12}(0,0) + \tfrac{1}{2}x_2^2 f_{22}(0,0)$$
$$= 6 + 3x_1^2 - 2x_1x_2 + 3x_2^2$$
$$= 6 + \tfrac{1}{3}(3x_1 - x_2)^2 + \tfrac{8}{3}x_2^2$$

Thus the normal form for the matrix of the quadratic form is

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and the stationary point is a local minimum.

LINEAR MATHEMATICS

335 01105 5